

STIC-Biotech/ChemLib

73147

From: Li, Ruixiang  
Sent: Tuesday, August 13, 2002 11:13 AM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search of Application NO:09/898,586

Please do a standard search on SEQ ID NO: 24 against both commercial and interference nucleic acid databases.

Thank you very much!

Ruixiang Li  
GAU 1646  
CM1 10E18  
Mail Box 10C01  
306-0282

(1)

10/13/02  
10/13/02  
10/13/02  
10/13/02

Point of Contact  
P. Sheppard  
Telephone number: (703) 308-4499

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: 8/23/02  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_



OM of: US-09-898-586-24 to: GenEmbl.\* out\_format : pfs

Date: Aug 21, 2002 8:18 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

#### Command line parameters:

```
-MODEL-frame+pn.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09898586/runat_16082002_205735_20549/app_query.fasta_1.373
-DB=GenEmbl -QFMT=fastap -SUFFIX=type -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEL=0.000 -LOOPEXT=0.000 -QGAPOP=4.500
-QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -RGAPOP=6.000
-RGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEX=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09898586@cgn1_1_4134
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1
```

#### Search information block:

```
Query: US-09-898-586-24
Query length: 310
Database: GenEmbl.*
Database sequences: 1797656
Search time (sec): 1826.930000
```

#### score\_list:

Sequence	Strg Orig	ZScore	EScore	Len	Documentation
gb_pat:AX078373	+ 1609.00	2439.04	1.7e-127	974	AX078373 Sequence 41 from Paten
gb_pat:AX197456	+ 1609.00	2438.73	1.8e-127	1012	AX197456 Sequence 21 from Paten
gb_pr:AC004889	- 1609.00	2398.40	3.1e-125	130030	AC004889 Homo sapiens clone
gb_pat:AX242191	+ 1602.00	2428.79	6.4e-127	930	AX242191 Sequence 939 from Paten
gb_pat:AX241442	+ 1584.00	2401.45	2.1e-125	930	AX241442 Sequence 190 from Paten
gb_pat:AX241473	+ 1584.00	2401.45	2.1e-125	930	AX241473 Sequence 221 from Paten
gb_pat:AX197438	+ 1584.00	2400.52	2.4e-125	1040	AX197438 Sequence 3 from Paten
gb_pat:AX197469	+ 1584.00	2400.52	2.4e-125	1040	AX197469 Sequence 34 from Paten
gb_pr:AL135904	- 1584.00	2366.96	1.8e-123	59130	AL135904 Human DNA sequence
gb_pr:AC005587	+ 1584.00	2359.75	4.5e-123	140915	AC005587 Homo sapiens PAC cl
gb_pat:AX197458	+ 1568.00	2376.43	5.3e-124	1014	AX197458 Sequence 23 from Paten
gb_ro:AX073444	+ 1330.00	2015.57	6.6e-104	933	AX073444 Mus musculus olfactory
gb_pr:AC091768	+ 1265.00	1877.60	3.2e-96	104955	AC091768 Homo sapiens BAC cl
gb_ro:AX073100	+ 1246.00	1887.96	8.5e-97	933	AX073100 Mus musculus olfactory
gb_pr:AC074386	+ 1138.00	1688.45	1.1e-85	66610	AC074386 Homo sapiens BAC cl
gb_pr:AF399598	+ 1129.00	1713.23	4.6e-87	650	AF399598 Homo sapiens clone OR2
gb_pat:AX242188	+ 1126.00	1705.70	1.2e-86	930	AX242188 Sequence 936 from Paten
gb_pat:AX242190	+ 1124.00	1702.11	1.9e-86	933	AX242190 Sequence 938 from Paten
gb_pat:AX078369	+ 1123.00	1695.69	4.3e-86	1792	AX078369 Sequence 37 from Paten
gb_ro:AX073723	+ 1067.00	1615.75	1.2e-81	966	AX073723 Mus musculus olfactory
gb_pat:AR092423	+ 1048.50	1582.89	8.3e-80	1713	AR092423 Sequence 1 from Paten
gb_ro:AX073724	+ 1046.00	1583.85	7.4e-80	966	AX073724 Mus musculus olfactory
gb_ro:AX073443	+ 1040.00	1574.95	2.3e-79	942	AX073443 Mus musculus olfactory
gb_ro:AX073820	+ 1002.00	1517.30	3.8e-76	933	AX073820 Mus musculus olfactory
gb_pat:AX241995	+ 992.00	1502.03	2.7e-75	942	AX241995 Sequence 743 from Paten
gb_htg:AC076959	+ 991.00	1457.74	7.8e-73	162320	AC076959 Homo sapiens chrom
gb_ro:AX073798	+ 990.00	1499.12	3.9e-75	927	AX073798 Mus musculus olfactory
gb_ro:AX073440	+ 975.00	1476.28	7.2e-74	933	AX073440 Mus musculus olfactory
gb_ro:AX073441	+ 974.00	1474.76	8.8e-74	933	AX073441 Mus musculus olfactory
gb_ro:AX073442	+ 973.00	1473.24	1.1e-73	933	AX073442 Mus musculus olfactory
gb_ro:MUSODORECA	+ 965.00	1463.88	3.5e-73	667	L14566 Mouse (K7) odorant recep
gb_htg:AC076959	- 947.50	1391.66	3.7e-69	162320	AC076959 Homo sapiens chrom
gb_pat:AX241886	+ 946.00	1431.75	2.2e-71	988	AX241886 Sequence 634 from Paten
gb_ro:AX073797	+ 943.00	1427.67	3.7e-71	933	AX073797 Mus musculus olfactory
gb_pat:AX242170	+ 920.00	1392.70	3.3e-67	937	AX242170 Sequence 918 from Paten
gb_pr:AC004889	+ 920.00	1351.73	6.3e-67	130030	AC004889 Homo sapiens clone
gb_pat:AX241944	+ 881.00	1333.38	6.6e-66	945	AX241944 Sequence 692 from Paten
gb_pat:AX242035	+ 854.00	1296.48	7.5e-64	576	AX242035 Sequence 783 from Paten
gb_pat:AX241845	+ 852.50	1290.04	1.7e-63	950	AX241845 Sequence 593 from Paten
gb_ro:AY073997	+ 851.00	1288.10	2.2e-63	912	AY073997 Mus musculus olfactory

```
gb_htg:AC090838 - 840.00 1250.72 2.6e-61 10985 ! AC090838 Homo sapiens chr
gb_htg:AC106354 - 838.50 1226.82 5.7e-60 148407 ! AC106354 Rattus norvegicu
gb_htg:AC106354 + 832.00 1216.95 2.0e-59 148407 ! AC106354 Rattus norvegicu
gb_pat:AX241887 + 813.00 1230.12 3.7e-60 941 ! AX241887 Sequence 635 from
gb_htg:AC068994_3 - 796.50 1166.97 1.2e-56 92197 ! Continuation (4 of 4) of
```

seq\_name: gb\_pat:AX078373

seq\_documentation\_block:

LOCUS AX078373 974 bp DNA linear PAT 22-FEB-2001

DEFINITION Sequence 41 from Patent WO0107612.

ACCESSION AX078373

VERSION AX078373.1 GI:13158042

KEYWORDS

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 974)

Au-Young,J., Bandman,O., Tang,Y.T., Yue,H., Azimzal,Y., Burford,N.,

Baughn,M.R., Lu,D.A., Hillman,J.L., Patterson,C. and Lal,P.

Receptors and associated proteins

Patent: WO 0107612-A 41 01-FEB-2001;

Incyte Genomics, Inc. (US)

location/Qualifiers

1..974

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/note="Incyte ID No: 2279267CBI"

BASE COUNT 206 a 270 c 214 g 284 t

ORIGIN

alignment\_scores:

Quality: 1609.00 Length: 310

Ratio: 5.190 Gaps: 0

Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-898-586-24 x AX078373 ..

Align seg 1/1 to: AX078373 from: 1 to: 974

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17

1 ATGGGGGACAAATATACATCCATCAGAGAGTTCTCTCTACTGGGATTTC 50

17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPhe 34

1 CGTTGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCTCCCTGTTCT 100

34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50

101 ACGTCTTCACCCCTGCTGGGGAACGGACCATCTGCGGCTCATCTCACTG 150

51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAla 67

151 GACTCCAGACTGCAGCCGCCCATGTACTTCTCTCTCCTCAGACCTGGCGGT 200

67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84

201 CGTCGACATCGCCCTACGCTGCAACACAGGTGCCCGGATGCTGTTGAACC 250

84 euleuHisProAlaIleProIleSerPheAlaGlyArgMetMetGlnThr 100

251 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGACAGACC 300

101 PheLeuPheSerThrPheAlaValThrIleCysLeuLeuValValMet 117

301 TTTCTGTTTTCACATTTTGTCTGTCACAGATGTCTCCCTGCTGTTGAT 350

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 134

```
351 GTCCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCA 400
134 lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150
401 TCATGACCTGGAGAGTGTGCATCACCTCGCGGTGACTTCTCGGACCACT 450
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuPro 167
451 GGAGTCCTTTATCTCTGATTCATCTTGTGTACTTCTACCTTTACCTT 500
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184
501 CTGTAGGCCCCAGAAAATTATCACTTTTGTGAAATCTTGGCTGTTC 550
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
551 TCAAACTTGCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 600
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerTy 217
601 GGAGCAATTTCTGGGCTGTGGGACCTTGTCCACAATTTGTAGTTCCATA 650
217 rmetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
651 TATGTGCATCTCTGTGCTATCTCTCAGATCCAAATCAAGGAAGTTCCAG 700
234 rglYsAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
701 GGAAGCCCTTGTGCACCTGTCTTCCACCTGTGTGTGATTTGGACTCTTT 750
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
751 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGAAACCCAA 800
267 sglGlnLysIleTyrLeuLeuPheHisSerLeuPheAsnProMetL 284
801 GGAGCAGAGAAATATCTCTGTGTTCACAGCCTCTTAAATCCCATGC 850
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
851 TCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAACTGAGAATACTTTG 900
301 LysArgValLeuGlyValGluArgAlaLeu 310
901 AAGAGAGTGTGGGAGTAGAAAGGCTTTA 930

seq_name: gb_pat:AX197456

seq_documentation_block:
LOCUS AX197456 1012 bp DNA linear PAT 29-AUG-2001
DEFINITION Sequence 21 from Patent WO0151632.
ACCESSION AX197456
VERSION AX197456.1 GI:15387837
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1012)
AUTHORS Padigaru,M., Prayaga,S.K., Tauplier,R.J., Mishra,V., Tchernev,V.T.,
Spytek,K.A. and Li,L.
TITLE Odorant receptor polypeptides and nucleic acids encoding same
JOURNAL Patent: WO 0151632-A 21 19-JUL-2001;
Curagen Corporation (US)
FEATURES
Source location/Qualifiers
1..1012
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 219 a 281 c 215 g 297 t
ORIGIN
alignment_scores:
```

```
Quality: 1609.00 Length: 310
Ratio: 5.190 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-898-586-24 x AX197456 ..
Align seg 1/1 to: AX197456 from: 1 to: 1012

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||
54 ATGGGAGACAAATATAACATCCATCACAGAGTTCCTCTACTGGGATTCC 103
17 ovalGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
|||||
104 CGTTGGCCCAAGGATTCAGATGCTCTCTTGGGCTCTTCTCCCTGTTCT 153
34 YrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
|||||
154 ACGTCTTACCCCTGCTGGGAGACGGGACCATACTGGGCTCATCTCACTG 203
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
|||||
204 GACTCCAGACTGCACGCCCCCATGTACTTCTTCTCTCACACCTGGCGGT 253
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
|||||
254 CGTCGACATCGCCTACGCGTCGACACAGGTGCCCGGATGCTGTCAACC 303
84 eulEuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
304 TCCGTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGCAGACC 353
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValIleMe 117
|||||
354 TTCTGTTCCTCCACTTTTGTGCTGCACAGAAATGTCCTCTGCTGGTGTAT 403
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
|||||
404 GTCCTATGATGTGTACGTGGGCAATCTGCCACCCCTCCGATATTGGCCA 453
134 lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
|||||
454 TCATGACCTGGAGAGTCTGCATCACCTCGCGGTGACTTCTCTGACCACT 503
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuPro 167
504 GGAGTCTTTATCTCTGATTCATCTTGTGTACTTCTTACCTTTACCTT 553
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
|||||
554 CTGTAGGCCCCAGAAAATTATCACTTTTGTGAAATCTTGGCTGTTC 603
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||||
604 TCAAACTTGCTGTGCAGATACCCACATCAATGAGAACATGGTCTTGGCC 653
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerTy 217
|||||
654 GGAGCAATTTCTGGGCTGTGGGACCTTGTCCACAATTTGTAGTTCCATA 703
217 rmetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
|||||
704 TATGTGCATCTCTGTGCTATCTTCAATCCAAATCAAGGAAGTTCCAGA 753
234 rglYsAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
|||||
754 GGAAGCCCTTGTGCACCTGTCTTCCACACCTGTGTGTGATTTGGACTCTTT 803
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
|||||
804 TATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGAAACCCCAA 853
```

```
267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnPrometL 284
|||||
854 GGAGCAGAGAATAATCTCTGCTGTTTACAGCCCTTTAATCCCATGC 903
|||||
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
|||||
904 TCAATCCCCCTTATCTGTAGTCTTAGAACTCAGAACTGAAGATACTTGTG 953
|||||
301 LysArgValLeuGlyValGluArgAlaLeu 310
|||||
954 AAGAGAGTGTGGGAGTAGAAGGCGTTTA 983
|||||

seq_name: gb_pr:AC004889

seq_documentation_block:
LOCUS AC004889 130030 bp DNA linear PRI 14-JAN-1999
DEFINITION Homo sapiens clone DJ0798C17, complete sequence.
ACCESSION AC004889
VERSION AC004889.1 GI:4156187
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 130030)
AUTHORS Waterston, R.H.
JOURNAL Unpublished
TITLE The sequence of Homo sapiens clone
REFERENCE 2 (bases 1 to 130030)
AUTHORS Waterston, R.H.
JOURNAL Direct Submission
TITLE Submitted (12-JUN-1998) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 3 (bases 1 to 130030)
AUTHORS Waterston, R.H.
JOURNAL Direct Submission
TITLE Submitted (14-JAN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jan 14, 1999 this sequence version replaced gi:3213108.
FEATURES
source
1. 130030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DJ0798C17"
BASE COUNT 38117 a 25410 c 26567 g 39936 t
ORIGIN

alignment_scores:
Quality: 1609.00 Length: 310
Ratio: 5.190 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
US-09-898-24 x AC004889/rev ..
Align seg 1/1 to reverse of: AC004889 from: 1 to: 130030

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPheP 17
|||||
111193 CGTGGCCCCAAGATTCAGATGCTCCTCTTGGGCTCTTCTCCCTGTTCT 111144
|||||
111243 ATGGGGACAATATATACATTCATCACAGAGTTCTCTACTGGGATTTC 111194
|||||
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
|||||
111193 CGTGGCCCCAAGATTCAGATGCTCCTCTTGGGCTCTTCTCCCTGTTCT 111144
|||||
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
|||||
111143 ACGTCTTCACCCCTGCTGGGGGAACGGGACCATACTGGGGCTCATCTCACTG 111094
|||||
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVal 67
```

```
|||||
111093 GACTCCAGACTGCACGCCCCCATGTACTTCTCTCCCTCACACCTGGCGGT 111044
|||||
67 lValAspIleAlaIleTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
|||||
111043 CGTCGACATCGCCCTACGCTGCAACACGGTGCCCGGATGCTGTGAACC 110994
|||||
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
110993 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGCAGACC 110944
|||||
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117
|||||
110943 TTTCCTGTTTCCACTTTTGGCTGTCAAGAAATGTCCTCCCTGGTGTGAT 110894
|||||
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
|||||
110893 GTCCATATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 110844
|||||
134 lMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
|||||
110843 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCTGGACCACT 110794
|||||
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProp 167
|||||
110793 GGAGTCTCTTATCTTGAATTCATCTGTGTACTTCTTACCTTTACCTTT 110744
|||||
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184
|||||
110743 CTGTAGGCCCCCAAGAAATTATACACTTTTGTGTGAATCTTGCGCTGTT 110694
|||||
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||||
110693 TCAAACTTGCCCTGTGCAGATACCCACATCAATGAGAACATGGCTTGCGC 110644
|||||
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
|||||
110643 GGAGCAATTTCTGGGCTGGTGGGACCCCTTGCCACAATGTAGTTGCATA 110594
|||||
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
|||||
110593 TATGTGCATCCCTGTGCTATCTTACATTCAGATTCAGAGGAAGTTCAGA 110544
|||||
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
|||||
110543 GGAAGCCTTCTGCACCTGCTTCTCCACCTCTGTGTGATGTGACTCTTT 110494
|||||
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
|||||
110493 TATGGCACAGCCATATATCATGTATGTGACCCAGATATGGGAACCCCAA 110444
|||||
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPrometL 284
|||||
110443 GGAGCAGAGAATAATCTCTGCTGTTTACAGCCCTTTAATCCCATGC 110394
|||||
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
|||||
110393 TCAATCCCCCTTATCTGTAGTCTTAGAACTCAGAACTGAAGATACTTGTG 110344
|||||
301 LysArgValLeuGlyValGluArgAlaLeu 310
|||||
110343 AAGAGAGTGTGGGAGTAGAAGGCGTTTA 110314
|||||

seq_name: gb_pat:AX242191

seq_documentation_block:
LOCUS AX242191 930 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 939 from Patent WO0127158.
ACCESSION AX242191
VERSION AX242191.1 GI:15799066
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
```

artificial sequence.  
REFERENCE 1 (bases 1 to 930)  
AUTHORS Bellenson, J., Smith, D., Lancet, D., Glusman, G., Fuchs, T. and Yanai, I.  
TITLE Olfactory receptor sequences  
JOURNAL Patent: WO 0127158-A 939 19-APR-2001;  
FEATURES Disjuncts (US) : YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)  
          Location/Qualifiers  
          1..930  
          /organism="synthetic construct"  
          /db\_xref="taxon:32630"  
          /note="(H38g789 nucleotide)"  
BASE COUNT 192 a 266 c 201 g 271 t  
ORIGIN  
alignment\_scores:  
          Quality: 1602.00 Length: 310  
          Ratio: 5.184 Gaps: 0  
Percent Similarity: 99.677 Percent Identity: 99.677  
alignment\_block:  
US-09-898-586-24 x AX242191 ..  
Align seg 1/1 to: AX242191 from: 1 to: 930  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
1 ATGGGGAGCAATATACATCATCACAGAGTCCCTCTGAGATTTC 50  
17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPhe 34  
1 CGTTGGCCCAAGATTACAGATGCTCTTGGGCTCTTCCCTGTTCT 100  
34 yRValPheThrLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
1 ACGTCTTCACCTCTGCTGGGGAACGGACCATCTGCGGCTCATCTCA 150  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAla 67  
151 GACTCCAGACTGCACGCCCATGACTCTCTCTCACACCTGGCGGT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84  
201 GCTGCACATCGCTACGCCCTCAACACAGGTGCCCGGATGCTGGAACC 250  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
251 TCCTGCATCCAGCCCAAGCCCATCTCTTTCGGGGCCGCATGATGACACC 300  
101 PheLeuPheSerThrPheAlaValThrGlyCysLeuLeuValValMe 117  
301 TTTCTGTTTCCACTTTGCTGTCACAGAAATGCTCTCTGCTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 134  
351 GTCCATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 400  
134 lMetThrTyrPArgValCysIleThrLeuAlaValThrSerTyrThr 150  
401 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCCCTGACCACT 450  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuPro 167  
451 GGAGTCCCTTTATCTGTATCATCTGTGTACTTCTACCTTACCCCTT 500  
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
501 CTGTAGGCCCCAGAAAATTATCACTTTTGTGAATCTGTGCTGTTC 550  
184 eulysLeuAlaCysAlaAspThrHisIleasnGluAsnMetValLeuAla 200  
551 TCAAACTTGCTGTGACAGATACCCACATCAATGAGAACATGTGTGGCC 600

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerTy 217  
601 GGAGCAATTTCTGGCTGTGGGACCCCTTGTCCACAATTGTACTTCATA 650  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln 234  
651 TATGTGCAATCTCTGTGTCTACTCTTCAGATCCATCAAGGGAAGTTCAG 700  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
701 GGAAGCCCTTCGACCTGCTCTCCACCTCTGTGTGATGGACTCTT 750  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnPro 267  
751 TATGGCAGACCATTAATGATGATGTGACCCAGATATGGAAACCCCAA 800  
267 sGluGlnLysLysTyrIleLeuLeuPheHisSerLeuPheAsnProMet 284  
801 GGAGCAGAGAAATATCTCCCGCTGTTCACAGCCCTTATATCCCATG 850  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
851 TCAATCCCTTATCTGTAGTCTTAGAAGCTCAGAAAGTGAAGAAATACTTG 900  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
901 AAGAGAGTCTGGAGATGAGAAAGGCTTTA 930  
seq\_name: gb\_pat:AX241442  
seq\_documentation\_block:  
LOCUS AX241442 930 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 190 from Patent WO0127158.  
ACCESSION AX241442  
VERSION AX241442.1 GI:15798317  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 930)  
AUTHORS Bellenson, J., Smith, D., Lancet, D., Glusman, G., Fuchs, T. and Yanai, I.  
TITLE Olfactory receptor sequences  
JOURNAL Patent: WO 0127158-A 190 19-APR-2001;  
FEATURES Disjuncts (US) : YEDA RESEARCH AND DEVELOPMENT COMPANY, LTD. (IL)  
          Location/Qualifiers  
          1..930  
          /organism="synthetic construct"  
          /db\_xref="taxon:32630"  
          /note="(H38g38 nucleotide)"  
BASE COUNT 193 a 265 c 202 g 270 t  
ORIGIN  
alignment\_scores:  
          Quality: 1584.00 Length: 310  
          Ratio: 5.160 Gaps: 0  
Percent Similarity: 99.032 Percent Identity: 99.032  
alignment\_block:  
US-09-898-586-24 x AX241442 ..  
Align seg 1/1 to: AX241442 from: 1 to: 930  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
1 ATGGGAGACAAATATACATCATCACAGAGATTCTCTCTACTGGGATTTC 50  
17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPhe 34  
51 CGTTGGCCCAAGATTACAGATGCTCTTGGGCTCTTCTCCCTGTTCT 100



34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
101 ACGTCTTCACCCCTGCTGGGGAACGGAGCACTACTGGGCTCATCTCACTG 150  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
151 GACTCCAGACTGCACGCCCCCATGTACTCTCTCTCTCACACCTGGCGGT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
201 CGTCGACATCGCCTACGCTGCAACAGCGTGGCCCCGAGATGCTGGAACC 250  
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
251 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGCAGACC 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
301 TTCTGTCTTCCACTTTGCTGTCACAGAAATGTCCTCCTCGTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
351 GTCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
401 TCATGACCTGGAGAGTCTGCATCACCCCTCGGCTGACTTCTTGACCACT 450  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167  
451 GGAGTCTTTTATCTTGATTCATCTTGTTACTTCTTACCTTTACCCCTT 500  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184  
501 CTGTAGGCCCCAGAAATTATCACCCTTTTGTGAAATCTTGGCTGTTC 550  
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
551 TCAAACTTGCTGTGCAGATACCCACATCATGAGAACATGTCTTGGCC 600  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
601 GGAGCAATTTCTGGGCTGTGGGAGCCCTGTGCCACAATGTGATTTTCATA 650  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
651 TATGTGCATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 700  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
701 GGAAGCCTTCCGACACCTGCTCTCCACCTCTGTGATGTGACTCGTT 750  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
751 TATGGCAGCCATTTATCATGTATGTTGGACCCAGATATGGAAACCCCAA 800  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPrometL 284  
801 GGAGCAGAGAAATATCTCTGCTGTTTCACAGCCTCTTTAAATCCCATGC 850  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
851 TCAATCCCTTATCTGTACTCTTAGAAGTCAAGAGTGAAGAAATACTTTG 900  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
901 AAGAGAGTGTGGGAGTGAAGAAAGGCTTTA 930

seq\_name: gb\_pat:AX241473

seq\_documentation\_block:

LOCUS AX241473 930 bp DNA linear PAT 26-SEP-2001  
DEFINITION Sequence 221 from Patent WO0127158.

ACCESSION AX241473  
VERSION AX241473.1 GI:15798348  
KEYWORDS  
SOURCE  
ORGANISM synthetic construct.  
REFERENCE 1 (bases 1 to 930)  
AUTHORS Bellenson,J., Smith,D., Lancelot,D., Glusman,G., Fuchs,T. and Yanal,I.  
TITLE Olfactory receptor sequences  
JOURNAL Patent: WO 0127158-A 221 19-APR-2001;  
FEATURES  
source location/Qualifiers  
1. 930  
/organism="synthetic construct"  
/db\_xref="taxon:32630"  
/note="(H38970 nucleotide)"  
BASE COUNT 193 a 265 c 202 g 270 t  
ORIGIN  
alignment\_scores:  
Quality: 1584.00 Length: 310  
Ratio: 5.160 Gaps: 0  
Percent similarity: 99.032 Percent identity: 99.032  
alignment\_block:  
US-09-898-24 x AX241473 ..  
Align seg 1/1 to: AX241473 from: 1 to: 930  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
1 ATGGGAGACAATATAMACATCATCAGAGACTTCCCTACTGGGATTGCC 50  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
51 CGTTGCCCAAGGATTTCAGATGCTCCTCTTGGGCTCTTCCCTGTTCT 100  
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
101 ACGTCTTCACCCCTGCTGGGGAACGGAGCCATACCTGGGCTCATCTCACTG 150  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
151 GACTCCAGACTGCACGCCCCCATGTACTTCTCTCTCACACCTGGCGGT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
201 CGTCGACATCGCCTACGCTGCAACAGCGTGGCCCCGAGATGCTGGAACC 250  
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
251 TCCTGCATCCAGCCCAAGCCCATCTCCTTGGCGGCGCATGATGCAGACC 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
301 TTCTGTCTTCCACTTTGCTGTCACAGAAATGTCCTCCTCGTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
351 GTCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
401 TCATGACCTGGAGAGTCTGCATCACCCCTCGGCTGACTTCCGGAGCACT 450  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167  
451 GGAGTCTTTTATCTTGATTCATCTTGTTACTTCTTACCTTTACCCCTT 500  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184

```
501 CTGTAGCCCCAGAAATTTATCACTTTTGTGAAATCTTGCTGTTTC 550
184 eulysleuAlaCysAlaAspThrHisIleAsnGluAsnMetValIleuAla 200
|||||
551 TCAAACTTGCTGTGCAGATACCCACATCAATGAGAAACATGTCCTGGCC 600
201 GAlaAlaIleSerGlyIleuValGlyProIleuSerThrIleValIleuValSerTy 217
|||||
601 GGAGCAATTTCTGGGCTGTGGGACCTTGTCACAAATTTGATGTTTCATA 650
217 rMetCysIleIleuCysAlaIleIleuGlnIleGlnSerArgGluValGlnA 234
|||||
651 TATGTGCATCCTCTGTGCTATCCTTCAGATCCAAATCAAGGAGAGTTCCAGA 700
234 rGlyAlaIlePheCysThrCysPheSerHisIleuCysValIleGlyIleuPhe 250
|||||
701 GGAAGCCCTCCGACACCTGCTCTCCACCTCTGTGATGTGACCTCGTT 750
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
|||||
751 TATGGCACAGCCATTATCATGTATGTGGACCCAGATATGGAAACCCCAA 800
267 sGluGlnIleLysIleTyrIleuIleuPheHisSerIleuPheAsnProMetL 284
|||||
801 GGAGCAGAAAGAAATATCTCCTGCTGTTCACAGCCTCTTAATCCCATGC 850
284 euAsnProIleuIleCysSerIleuArgAsnSerGluValLysAsnThrIleu 300
|||||
851 TCAATCCCCCTTATCTGTAGCTTAGGAACCTCAGAAAGTGAAGAAATACTTTG 900
301 LysArgValIleuGlyValGluArgAlaIleu 310
|||||
901 AAGAGAGTGTGGGAGTAGAAGGCTTTA 930
```

seq\_name: gb\_pat:AX197438

seq\_documentation\_block:

LOCUS AX197438 1040 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 3 from Patent WO0151632.  
ACCESSION AX197438  
VERSION AX197438.1 GI:15387828

KEYWORDS

SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 1040)

AUTHORS Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T.,  
Spytek,K.A. and Li,L.

TITLE Odorant receptor polypeptides and nucleic acids encoding same

JOURNAL Patent: WO 0151632-A 3 19-JUL-2001;  
Curagen Corporation (US)

FEATURES location/Qualifiers  
source 1..1040  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 229 a 285 c 221 g 305 t

ORIGIN

alignment\_scores:  
Quality: 1584.00 Length: 310  
Ratio: 5.160 Gaps: 0  
Percent Similarity: 99.032 Percent Identity: 99.032

alignment\_block:

US-09-898-586-24 x AX197438 ..

Align seg 1/1 to: AX197438 from: 1 to: 1040

```
1 MetGlyAspAsnIleThrSerIleThrGluPheIleuIleuGlyPhePr 17
|||||
82 ATGGAGACATATATACATCCATCAAGAGAGTTCCTCTACTGGGATTTCC 131
```

```
17 oValGlyProArgIleGlnMetIleuIleuPheGlyIleuPheSerIleuPheThr 34
|||||
132 CGTTGGCCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCTCCCTGTTCT 181
34 yrValPheThrIleuIleuGlyAsnGlyThrIleIleuGlyIleuIleSerIleu 50
|||||
182 ACGTCTTACCCCTGCTGGGGAAGGAGACCATATCTGGGGCTCATCTCACTG 231
51 AspSerArgIleuHisAlaProMetTyrPhePheIleuSerHisIleuAlaIle 67
|||||
232 GACTCCAGACTGCACGCCCCCATGTACTTCTTCTCTCACACCTGGCGGT 281
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetIleuValAsnL 84
|||||
282 CGTCGACATCGCCTACGCGCTGCACACAGGTTGCCCGGATGCTGTGAACC 331
84 euleuHisProAlaIleLysProIleSerPheAlaGlyArgMetMetClnThr 100
|||||
332 TCCTGCATCCAGCCCAAGCCCATCTCCTTGGCGGGCGGCATGATGCAGACC 381
101 PheIleuPheSerThrPheAlaValThrGluCysIleuIleuValIleuValMe 117
|||||
382 TTTCTGTGTTCACACTTTTGTGCTGCACAGAATGTCTCTCTGTTGATGAT 431
117 tSerTyrAspIleuTyrValAlaIleCysHisProIleuArgTyrIleuAlaI 134
|||||
432 GTCTATGATCTGTACGTGGCCATCTGCACACCCCTCCGATATTTGGCCA 481
134 leMetThrThrPargValCysIleThrIleuAlaValThrSerTrpThrThr 150
|||||
482 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCTCGACCACT 531
151 GlyValIleuSerIleuIleHisIleuValIleuIleuProIleuProPh 167
|||||
532 GGAGTCCTTTATCCTTGATTCATCTTGTGTACTTCTTACCTTTACCCCTT 581
167 eCysArgProGlnIleTyrHisPhePheCysGluIleIleuAlaValL 184
|||||
582 CTGTAGGCCCCAGAAATTTATCACTTTTGTGGAATCTTGCTGTTC 631
184 eulysleuAlaCysAlaAspThrHisIleAsnGluAsnMetValIleuAla 200
|||||
632 TCAAACTTGCTGTGCAGATACCCACATCAATGAGAACATGTCCTTGGCC 681
201 GAlaAlaIleSerGlyIleuValGlyProIleuSerThrIleValIleuValSerTy 217
|||||
682 GGAGCAATTTCTGGGCTGTGGGACCTTGTCACAAATTTGATGTTTCATA 731
217 rMetCysIleIleuCysAlaIleIleuGlnIleGlnSerArgGluValGlnA 234
|||||
732 TATGTGCATCCTCTGTGCTATCTTCAGATCCAAATCAAGGAGAGTTCCAGA 781
234 rGlyAlaIlePheCysThrCysPheSerHisIleuCysValIleGlyIleuPhe 250
|||||
782 GGAAGCCCTCCGACACCTGCTCTCCACCTCTGTGATGTGACCTCGTT 831
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
|||||
832 TATGGCACAGCCATTATCATGTATGTTGAGACCCAGATATGGAAACCCCAA 881
267 sGluGlnIleLysIleTyrIleuIleuPheHisSerIleuPheAsnProMetL 284
|||||
882 GGAGCAGAAAGAAATATCTCCTGCTGTTCACAGCCTCTTTAATCCCATGC 931
284 euAsnProIleuIleCysSerIleuArgAsnSerGluValLysAsnThrIleu 300
|||||
932 TCAATCCCCCTTATCTGTAGCTTTAGGAACCTCAGAAAGTGAAGAAATACTTTG 981
301 LysArgValIleuGlyValGluArgAlaIleu 310
|||||
982 AAGAGAGTGTGGGAGTAGAAGGCTTTA 1011
```



seq\_name: gb\_pat:AX197469

seq\_documentation\_block:

LOCUS AX197469 1040 bp DNA linear PAT 29-AUG-2001

DEFINITION Sequence 34 from Patent WO0151632.

ACCESSION AX197469

VERSION AX197469.1 GI:15387841

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1040)

AUTHORS Padigaru,M., Prayaga,S.K., Taupier,R.J., Mishra,V., Tchernev,V.T., Spytek,K.A. and Li,L.

TITLE Odorant receptor polypeptides and nucleic acids encoding same

JOURNAL Patent: WO 0151632-A 34 19-JUL-2001;

FEATURES

source 1.1040

location/Qualifiers

BASE COUNT 229 a 285 c 221 g 305 t

ORIGIN

alignment\_scores:

Quality: 1584.00 Length: 310

Ratio: 5.160 Gaps: 0

Percent Similarity: 99.032 Percent Identity: 99.032

alignment\_block:

US-09-898-586-24 x AX197469 ..

Align seg 1/1 to: AX197469 from: 1 to: 1040

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17

|||||

82 ATGGGAGACAATATACATCATCAGAGAGTTCCTCTACTGGGATTTC 131

17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34

|||||

132 CGTTGGCCCAAGGATTACAGATGCTCTCTTTGGGCTCTTCTCCCTGTCT 181

34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50

|||||

182 ACGTCTTCACCCCTGCTGGGGAACGGACCATACGGGGCTCATCTCACTG 231

51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67

|||||

232 GACTCCAGACTGCAGGCCCATGTACTCTCTCTCTCACACCTGCGCGT 281

67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84

|||||

282 CGTCGACATCGCCTACGCTGCAACACGGTGCCTCCGATGCTGTGAACC 331

84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100

|||||

332 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGACAGCC 381

101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValme 117

|||||

382 TTTCGTGTTTCCACTTTTGTGTCACAGAATGTCTCTCTGTTGTTGAT 431

117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrIleuAlaI 134

|||||

432 GTCCATAGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 481

134 leMetThrTTPArgValCysIleThrLeuAlaValThrSerTyrThrThr 150

|||||

482 TCATGACCTGGAGAGTCTGCATCACCCCTGCGGTGACTTCTTGACACACT 531

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167

|||||

532 GGAGTCTCTTTATCCCTGATCATCTTGTTACTTCTACCTTTACCTT 581

167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184

|||||

582 CTGTAGCCCCCAGAAAATTTATCATTCTTTTGTGAATCTTGCTGTTTC 631

184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200

|||||

632 TCAAACTTGCTGTGCAGATACCCACATCATGAGAACATGGTCTTGCC 681

201 GlValIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217

|||||

682 GGAGCAATTTCTGGGCTGGTGGACCCCTTGTCACCAATTTGATTTCATA 731

217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234

|||||

732 TATGTGCATCTCTGTGCTATCTTCAGATCCAAATCAAGGAAGTTCCAGA 781

234 rglYsAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250

|||||

782 GGAAGCCCTTCCGACCTGCTTCCACACCTGTGTGATTTGACACTCGTT 831

251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267

|||||

832 TATGGCAGACCATATATCATGATGTTGGACCCAGATATGGAAACCCCAA 881

267 sgluGlnLysIleTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284

|||||

882 GGAGCAGAGAATAATATCTCTGCTGTTTTCACAGCCTCTTAATCCCATGC 931

284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300

|||||

932 TCAATCCCTTATCTGTAGTCTTAGCAACTCAGAACTGAAGAATACCTTG 981

301 LysArgValLeuGlyValGluArgAlaLeu 310

|||||

982 AAGAGAGTGTGGAGTAGAAAGGCTTTA 1011

seq\_name: gb\_pr:AL135904

seq\_documentation\_block:

LOCUS AL135904 59130 bp DNA linear PRI 20-JUL-2000

DEFINITION Human DNA sequence from clone RP5-1005H11 on chromosome 6 Contains part of the gene for a 7 transmembrane receptor (rhodopsin family) (olfactory receptor like) protein, the gene for WUGSC:4.DJ0988G15.3 protein, part of the PDNP3 (phosphodiesterase I/nucleotide pyrophosphatase 3) gene, ESTs, STSS and GSSs, complete sequence.

ACCESSION AL135904

VERSION AL135904.11 GI:71593399

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 59130)

AUTHORS Lovell,J.

TITLE Direct Submission

JOURNAL Submitted (05-JUL-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

COMMENT

On Mar 6, 2000 this sequence version replaced gi:7105768.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure'

feature key.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; tr: TrEMBL; wp: WORMPEP; Information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr6>  
RP5-1005H11 is from the library RPCI-5 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see <http://bacpac.med.buffalo.edu/>  
VECTOR: pCYPAC2  
IMPORTANT: This sequence is not the entire insert of clone RP5-1005H11. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
The true right end of clone RP5-1005H11 is at 59130 in this sequence. The true right end of clone RP5-914M13 is at 105 in this sequence.

FEATURES	
source	Location/Qualifiers
repeat_region	1..59130 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="6" /clone="RP5-1005H11" /clone_11b="RPCI-5"
repeat_region	1..68 /note="AluS repeat: matches 1..68 of consensus"
repeat_region	546..844 /note="AluJo repeat: matches 2..301 of consensus"
repeat_region	870..940 /note="U4 repeat: matches 1..72 of consensus"
repeat_region	1691..2906 /note="L2 repeat: matches 1425..2737 of consensus"
repeat_region	3457..3960 /note="HAL1 repeat: matches 420..971 of consensus"
repeat_region	3971..4106 /note="MTFD repeat: matches 1..129 of consensus"
repeat_region	4249..4282 /note="17 copies 2 mer tg 100% conserved"
repeat_region	4966..5147 /note="AluJb repeat: matches 1..182 of consensus"
repeat_region	5216..5307 /note="L2 repeat: matches 2405..2503 of consensus"
repeat_region	5495..5794 /note="AluSx repeat: matches 1..296 of consensus"
repeat_region	5891..6180 /note="AluSx repeat: matches 1..291 of consensus"
repeat_region	7607..7618 /note="L1M1 repeat: matches 4856..4867 of consensus"
repeat_region	7619..8347 /note="L1PA16 repeat: matches 5421..6156 of consensus"
repeat_region	8348..8898 /note="L1M1 repeat: matches 4866..5403 of consensus"
repeat_region	9018..9866 /note="TIGGER1 repeat: matches 1..836 of consensus"
repeat_region	9867..10189 /note="AluSx repeat: matches 1..309 of consensus"
repeat_region	10190..11723 /note="TIGGER1 repeat: matches 836..2418 of consensus"
repeat_region	12404..12715 /note="AluSx repeat: matches 1..312 of consensus"
misc_feature	complement(13947..14176) /note="match: GSS: Em:B74942"
repeat_region	14280..14383 /note="AluSg/x repeat: matches 190..294 of consensus"
repeat_region	15340..15508 /note="L2 repeat: matches 2563..2750 of consensus"
repeat_region	15684..15968 /note="AluSx repeat: matches 35..312 of consensus"
repeat_region	16527..16612 /note="MIR repeat: matches 50..138 of consensus"
repeat_region	16731..17027 /note="AluSg repeat: matches 1..295 of consensus"
repeat_region	17609..17747 /note="MER91A repeat: matches 1..153 of consensus"
repeat_region	17867..18162 /note="AluY repeat: matches 1..299 of consensus"
repeat_region	18461..18950 /note="MER1A repeat: matches 2..527 of consensus"
repeat_region	19185..19224 /note="20 copies 2 mer tg 85% conserved"
misc_feature	19474..19712 /note="match: STS: Em:G16149"
repeat_region	20037..20361 /note="AluSx repeat: matches 1..311 of consensus"
gene	complement(21566..22384) /gene="dj1005H11.1"
CDS	complement(21566..22384) /gene="dj1005H11.1"> /note="match: proteins: Sw:P47881 Tr:O70265 Tr:O70266 Tr:O70267 Tr:O70268 Tr:O62942 Tr:O95918 Tr:O95047"
	/codon_start=1 /evidence="not_experimental"
	/product="dj1005H11.1 (7 TRANSMEMBRANE RECEPTOR (RHODOPSIN FAMILY) (OLFACTORY RECEPTOR LIKE) PROTEIN))"
	/protein_id="GAB99212.1"
	/db_xref="GI:9368991"
	/translation="LGGTILGLISDSRLHAPMYFELSHLAVVDIAYACNTVPRMLV NLHPAKPISFAGRMQTELESTFAVECLLVMSYDLYVAICHPLRYLAIMTRVC ITLAVTSWTTGVLSLIHVLVLLPLPCRPQKIYHFECELAVLKLACADTHINENMV LGAISGLVGPISTIVSYMCITCAIOISREVORAKFRFCESHLACVIGLYGTALII MYVGPRYGNPKQKKYLLFHSLEPNMNPILICSLRNSYKNTLKRVLGVERAL"
misc_feature	complement(22229..22521) /note="match: STS: Em:H39853"
repeat_region	22385..22489 /note="L2 repeat: matches 2461..2551 of consensus"
repeat_region	22696..23011 /note="AluSx repeat: matches 1..312 of consensus"
misc_feature	23044..23513 /note="match: STS: Em:A1066560"
misc_feature	complement(23123..23716) /note="match: GSS: Em:AQ276140"
misc_feature	complement(23300..23543) /note="match: GSS: Em:AQ058777"
misc_feature	23751..24305 /note="match: GSS: Em:AQ798220"
misc_feature	23753..24176 /note="match: GSS: Em:AQ550610"
misc_feature	23762..24465 /note="match: GSS: Em:B86674"
repeat_region	24220..24343 /note="FLAM_C repeat: matches 1..124 of consensus"
misc_feature	complement(24681..25395) /note="match: GSS: Em:AQ627093"
misc_feature	complement(24780..25383) /note="match: GSS: Em:AQ627184"
repeat_region	26053..26548 /note="TIGGER1 repeat: matches 1..521 of consensus"
repeat_region	26549..26861 /note="L1M4 repeat: matches 4025..4339 of consensus"
repeat_region	26865..27121 /note="Tigger3b repeat: matches 2..272 of consensus"
repeat_region	27122..27424 /note="AluSx repeat: matches 1..302 of consensus"
repeat_region	27425..27579 /note="Tigger3b repeat: matches 272..451 of consensus"
repeat_region	27582..27658 /note="Tigger3b repeat: matches 1155..1231 of consensus"
repeat_region	27659..29067 /note="L1M4 repeat: matches 2585..4035 of consensus"
repeat_region	29189..29287 /note="L1M4 repeat: matches 2499..2600 of consensus"
misc_feature	29316..29499 /note="match: STS: Em:AA931824"

misc\_feature 29507..29990  
/note="match: STS: Em:AA931824"  
misc\_feature 29530..29981  
/note="match: GSS: Em:AQ356766"  
misc\_feature 29530..29737  
/note="match: GSS: Em:AQ253729"  
misc\_feature 29533..29855  
/note="match: GSS: Em:AQ319835"  
misc\_feature 29538..29943  
/note="match: STS: Em:AA235111"  
misc\_feature 29560..29787  
/note="match: STS: Em:G27651 Em:T99793"  
gene complement(29779..32112)  
/gene="dj1005H1.2"  
CDS complement(29779..32112)  
/gene="dj1005H1.2"  
/db\_xref="GI:9368992"  
/translation="MEEPGATPPYLGVLVEELGRVVAALPESMRPDENPYGFPSELV"

alignment\_scores: Quality: 1584.00 Length: 310  
Ratio: 5.160 Gaps: 0  
Percent Similarity: 99.032 Percent Identity: 99.032

alignment\_block:  
us-09-898-586-24 x AL135904/rev ..

Align seg 1/1 to reverse of: AL135904 from: 1 to: 59130

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
|||||  
22498 ATGGAGACAAATATACATCCATCCAGAGAGTTCTCTCTGCGATTCC 22449  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
|||||  
22448 CGTGGCCCCAAGSATTCAAGATGCTCTTGGGCTCTCTCCCTGTTCT 22399  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
22398 ACGTCTCACCCCTGCTGGGAGACGGGACCATACTGGGGCTCATCTCACTG 22349  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67  
|||||  
22348 GACTCCAGACTGACGCCCCCATGTACTTCTCTCTCTCACACCTGGCGGT 22299  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
|||||  
22298 CGTCGACATCGCCTACGCGCTGCACACGGGTGCCCGGATGCTGTGAACC 22249  
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
22248 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGGCCGATGATGCAGACC 22199  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
22198 TTTCTGTTCACATTCTTGGCTGTGCACAGAATGTCCTCTGTTGGTGTAT 22149  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
22148 GTCCATATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 22099  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
|||||  
22098 TCATGACCTGGAGAGTCTGCATCACCTCGCGGTGACTTCCTGAGACCACT 22049  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167  
|||||  
22048 GGAGTCCTTTTATCCTGTATTCATTCATTGTGTACTTCTTACCTTTACCTTT 21999

167 eCysArgProGlnLysIleThrHisPhePheCysGluIleLeuAlaValL 184  
|||||  
21998 CTGTAGGCCCCAGAAAATTATCACCCTTTTGTGAAATCTGTGGCTGTTG 21949  
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
21948 TCAAACTTGCTGTGCAGATACCCACATCAATGAGAAATGTTCTGGCC 21899  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
|||||  
21898 GGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAATTGTAGTTGATA 21849  
217 tMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgLysValGlnA 234  
|||||  
21848 TATGTGCATCCTCTGTGCTATCCCTTCAGATCCAAATCAAGGAAGTCAGA 21799  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
21798 GGAAGCCTTCCGACACCTGCTCTCCACCTCTGTGTGATTGACTGCTT 21749  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
|||||  
21748 TATGGCACAGCATATATCATGTATGTGTGACCCAGATATGGAAACCCAA 21699  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPrometL 284  
|||||  
21698 GGAGCAGAGAATAATATCTCTGCTGCTGTTTTCACAGCCTCTTATATCCATGC 21649  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||||  
21648 TCAATCCCCTTATCTGTAGCTTAGAAGTCAAGTGAAGAATACTTTG 21599  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
21598 AAGAGAGTGTGAGAGTAGAAGGCTTTA 21569  
seq\_name: gb\_pr:AC005587  
seq\_documentation\_block:  
LOCUS AC005587 140915 bp DNA linear PRI 21-DEC-1999  
DEFINITION Homo sapiens PAC clone RP5-988615 from 7q33-q35, complete sequence.  
ACCESSION AC005587  
VERSION AC005587.1 GI:4156166  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS Hou,S., Wohlmann,P. and Le,T.  
TITLE The sequence of Homo sapiens PAC clone RP5-988615  
JOURNAL Unpublished  
REFERENCE  
AUTHORS 2 (bases 1 to 140915)  
TITLE Waterston,R.H.  
JOURNAL Direct Submission  
REFERENCE  
AUTHORS 3 (bases 1 to 140915)  
TITLE Waterston,R.H.  
JOURNAL Direct Submission  
REFERENCE  
AUTHORS 4 (bases 1 to 140915)  
TITLE Waterston,R.  
JOURNAL Direct Submission  
REFERENCE  
AUTHORS Submitted (15-JAN-1999) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
5 (bases 1 to 140915)  
TITLE Waterston,R.  
JOURNAL Direct Submission

JOURNAL

COMMENT

Submitted (21-DEC-1999) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Jan 14, 1999 this sequence version replaced gi:3907499.

----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc>  
Contact: [sapiens@watsn.wustl.edu](mailto:sapiens@watsn.wustl.edu)  
----- Summary Statistics  
Center project name: H\_DJ0988G15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCL-5, prepared by Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>) using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.

The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.

VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP5-988G15;  
actual end is at 140915 of RP5-988G15.

Location/Qualifiers

1. .140915

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="7"

/map="7q33-q35"

/clone="RP5-988G15"

/clone\_lib="RPCL-5"

13. .2090

/rpt\_family="L1"

2097. .2299

/rpt\_family="L1"

2300. .2603

/rpt\_family="Alu"

2604. .4120

/rpt\_family="L1"

4123. .4166

/rpt\_family="(TA)n"

4168. .4307

/rpt\_family="L1"

6635. .6668

/rpt\_family="AT\_rich"

6829. .7260

/rpt\_family="L1"

7261. .7642

/rpt\_family="MaLR"

repeat\_region

7643. .7923

repeat\_region

/rpt\_family="L1"

repeat\_region

7982. .8283

repeat\_region

/rpt\_family="Alu"

repeat\_region

8286. .8412

repeat\_region

/rpt\_family="L1"

repeat\_region

8413. .8708

repeat\_region

/rpt\_family="Alu"

repeat\_region

8709. .8838

repeat\_region

/rpt\_family="L1"

repeat\_region

8840. .8886

repeat\_region

/rpt\_family="MaLR"

repeat\_region

9111. .9139

repeat\_region

/rpt\_family="AT\_rich"

repeat\_region

9204. .9490

repeat\_region

/rpt\_family="Alu"

repeat\_region

9646. .9924

repeat\_region

/rpt\_family="Alu"

repeat\_region

10286. .10455

repeat\_region

/rpt\_family="L2"

repeat\_region

10759. .10838

repeat\_region

/rpt\_family="Mariner"

repeat\_region

11127. .11266

repeat\_region

/rpt\_family="L2"

repeat\_region

11850. .12154

repeat\_region

/rpt\_family="Alu"

repeat\_region

12172. .12211

repeat\_region

/rpt\_family="AT\_rich"

repeat\_region

12352. .12469

repeat\_region

/rpt\_family="(TACG)n"

repeat\_region

12458. .12607

repeat\_region

/rpt\_family="(CATA)n"

repeat\_region

12580. .12733

repeat\_region

/rpt\_family="(TA)n"

repeat\_region

12660. .12777

repeat\_region

/rpt\_family="(CATA)n"

repeat\_region

12778. .13031

repeat\_region

/rpt\_family="Alu"

repeat\_region

13226. .13535

repeat\_region

/rpt\_family="L2"

repeat\_region

15279. .15520

repeat\_region

/rpt\_family="MIR"

repeat\_region

15954. .16022

repeat\_region

/rpt\_family="MaLR"

repeat\_region

16023. .16332

repeat\_region

/rpt\_family="MER2\_type"

repeat\_region

16333. .16371

repeat\_region

/rpt\_family="MaLR"

repeat\_region

16630. .16742

repeat\_region

/rpt\_family="MaLR"

repeat\_region

16816. .17068

repeat\_region

/rpt\_family="L1"

repeat\_region

17088. .17348

repeat\_region

/rpt\_family="L1"

repeat\_region

17857. .17979

repeat\_region

/rpt\_family="Alu"

repeat\_region

18014. .18506

repeat\_region

/rpt\_family="MaLR"

repeat\_region

18707. .19081

repeat\_region

/rpt\_family="MaLR"

repeat\_region

19135. .20717

repeat\_region

/rpt\_family="Retroviral"

repeat\_region

20728. .20748

repeat\_region

/rpt\_family="AT\_rich"

repeat\_region

21444. .21599

repeat\_region

/rpt\_family="MIR"

misc\_feature

21688. .22279

repeat\_region

/note="60S ribosomal protein L15 pseudogene"

repeat\_region

22506. .22540

repeat\_region

/rpt\_family="AT\_rich"

repeat\_region

22557. .22583

repeat\_region

/rpt\_family="AT\_rich"

repeat\_region

23139. .23162



```
repeat_region /rpt_family="(TAA)n"
23696..24012 /rpt_family="Alu"
repeat_region /rpt_family="AT_rich"
24088..24128 /rpt_family="AT_rich"
repeat_region /rpt_family="AT_rich"
24827..24849 /rpt_family="AT_rich"
repeat_region /rpt_family="Alu"
24850..25147 /rpt_family="Alu"
repeat_region /rpt_family="AT_rich"
25148..25173 /rpt_family="AT_rich"
repeat_region /rpt_family="AT_rich"
25951..25974 /rpt_family="(CAAAA)n"
repeat_region /rpt_family="Alu"
26066..26144 /rpt_family="Alu"
repeat_region /rpt_family="L1"
26388..26653 /rpt_family="L1"
repeat_region /rpt_family="L2"
27148..27600 /rpt_family="L2"
repeat_region /rpt_family="Alu"
28002..28294 /rpt_family="Alu"
```

alignment\_scores:
Quality: 1584.00 Length: 310
Ratio: 5.160 Gaps: 0
Percent Similarity: 99.032 Percent Identity: 99.032

alignment\_block:
US-09-898-586-24 x AC005587 ..

Align seg 1/1 to: AC005587 from: 1 to: 140915

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
101599 ATGGGAGACATATATACATCATCAGAGAGTTCTCTCTACTGGGATTTC 101648
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
101649 CGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCTGTTCT 101698
34 yValPheThrLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
101699 ACGTCTTCACCCCTGCTGGGGAACGGGACCATCTGCGGCTCATCTCACTG 101748
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
101749 GACTCCAGACTGCACGCCCCCATGTACTTCTCTCTCTCACACCTGGCGGT 101798
67 lValAspIleAlaIleThrAlaCysAsnThrValProArgMetLeuValAsnL 84
101799 CGTCGACATCGCCCTACGCTTCACACACGCGGCGGATGCTGCTGAACC 101848
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
101849 TCCTGCATCCAGCCCAACCCCATCTCCTTGGCGGCGGCGCATGATGCAGACC 101898
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValMe 117
101899 TTTCTGTCTTCCACTTTTCTGCTGCACAGAAATGTCCTCCTGCTGTGTAT 101948
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
101949 GTCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 101998
134 leMetThrTyrPargValCysIleThrLeuAlaValThrSerTyrPheThr 150
101999 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCCGTGACCACT 102048
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
102049 GGAGTCCCTTTATCTTATCATCTGTGTACTTCTACCTTTACCTTT 102098
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
```

```
102099 CTGTAGGCCCCAGAAAATTTATCACTTTTTTTGTGAATCTTGGCTGTC 102148
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
102149 TCAAACTTGCTGTGCAGATACCACATCATGAGACATGCTGTGCC 102198
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleValSerTy 217
102199 GGAGCAATTTCTGGGCTGTGGGACCTTGTCCACAAATTGATTTCATA 102248
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
102249 TATGTGATCCTCTGTGCTATCCTTCAGATCCAATCAAGGAAGTTCAGA 102298
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
102299 GGAAGCCTTCCGACACCTGCTTCCACACCTCTGTGTGATTTGACTCGTT 102348
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
102349 TATGGCAGACCATTTATCATGTATGTGAGACCCAGATATGGGAACCCCAA 102398
267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL 284
102399 GGAGCAGAGAATAATATCTCCTGCTGTTTTCACAGCCCTTTAATCCCATGC 102448
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
102449 TCAATCCCTTATCTGTAGTCTTTAGGAAGTCAAGAGTGAAGAACTTTG 102498
301 LysArgValLeuGlyValGluArgAlaLeu 310
102499 AAGAGAGTGGCTGGAGAGTAGAAGGCGTTTA 102528
```

seq\_name: gb\_pat:AX197458

seq\_documentation\_block:

LOCUS AX197458 1014 bp DNA linear PAT 29-AUG-2001  
DEFINITION Sequence 23 from Patent WO0151632.  
ACCESSION AX197458  
VERSION AX197458.1 GI:15387838  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1014)  
AUTHORS Padigaru,M., Prayaga,S.K., Taupler,R.J., Mishra,V., Tchernev,V.T.,  
Spytek,K.A. and Li,L.  
TITLE Odorant receptor polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0151632-A 23 19-JUL-2001;  
Curagen Corporation (US)  
FEATURES  
source location/Qualifiers  
1..1014  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 218 a 281 c 216 g 299 t  
ORIGIN

alignment\_scores:
Quality: 1568.00 Length: 311
Ratio: 5.074 Gaps: 1
Percent Similarity: 99.357 Percent Identity: 99.357

alignment\_block:
US-09-898-586-24 x AX197458 ..

Align seg 1/1 to: AX197458 from: 1 to: 1014

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||
```



55 ATGGGGGACAAATATAACATCCATCACAGAGTTCCTCCTACTGGGATTTC 104  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34  
105 CGTTGGCCCAAGATTCAGATGCTCCTTTGGGGCTCTCTCCCTGTCT 154  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
155 ACGTCTTCACTGCTGGGGAACGGACCATACTGGGGCTCATCTCAGTC 204  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
205 GACTCCAGACTGCACGCCCTCTG. TACTTCTTCTCTCCACACCTGGCGGT 253  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
254 GCTGCACATCGCTACGCGCTGCACACGGTGCCTCCGATGCTGTGAAC 303  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
304 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGGCCGATGATGCAGACC 353  
101 pheLeuPheSerThrPheAlaValThrGlyCysLeuLeuValValMe 117  
354 TTTCTGTTCACACTTTTGGCTGTCAAGAAATGTCTCTCTGCTGTGAT 403  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
404 GTCCATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTTGCCA 453  
134 lMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
454 TCATGACCTGCAGAGTCTGCATCACCTCGCGGTGACTTCTGGACCACT 503  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProh 167  
504 GGAGTCTTTTATCTTGTGATTCATCTTGTGTTACTTCTTACCTTTACCCTT 553  
167 ecysArgProGlnLysIleTyrHis. PhePheCysGlyIleLeuAlaVal 183  
554 CTGTAGGCCCCAGAAATTTATCATTCTTTTGTGAATCTTGGCTGTT 603  
184 LeuLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAl 200  
604 CTCAAACTTGCTGTGCAGATACCCACATCAATGAGACATGCTCTGGC 653  
200 aGlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerT 217  
654 CGGAGCAATTTCTGGGCTGTGGGACCCCTTGTCCACAAATGTAGTTTCAT 703  
217 yMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln 233  
704 ATATGTGCATCCTCTGTGTCTATCTTCAGATCCAATCAAGGGAAGTTCAG 753  
234 ArgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPh 250  
754 AGGAAAGCCCTTCTGCACCTGCTTCCACACTCTGTGTGATTTGGACTCTT 803  
250 eTyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProL 267  
804 TTATGGCACAGCCATATCATGTATGTGGACCCAGATATGGGAACCCCA 853  
267 ysgLuglnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPromet 283  
854 AGGAGCAGAAATAATATCTCTGCTTTCACAGCCCTTTTAAATCCCATG 903  
284 LeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrIle 300  
904 CTCATATCCCTTATCTGTAGTCTTAGGAACTCAGAAAGTCAAGATACTTT 953  
300 uLysArgValLeuGlyValGluArgAlaLeu 310  
954 GAAGAGAGTGTGGAGTAGAAAGGCTTTA 984

seq\_name: gb\_ro:AY073444  
seq\_documentation\_block:  
LOCUS AY073444 933 bp DNA linear ROD 04-FEB-2002  
DEFINITION Mus musculus olfactory receptor MOR261-6 gene, complete cds.  
ACCESSION AY073444  
VERSION AY073444.1 GI:18480185  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Zhang, X. and Firestein, S.  
TITLE The olfactory receptor gene superfamily of the mouse  
JOURNAL Nat. Neurosci. 5 (2), 124-133 (2002)  
PUBMED 11802173  
REFERENCE  
AUTHORS Adams, M.  
TITLE Direct Submission  
JOURNAL Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
FEATURES  
source location/Qualifiers  
1..933  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="6"  
<1..>933  
/product="Olfactory receptor MOR261-6"  
1..933  
/codon\_start=1  
/evidence=not\_experimental  
/product="Olfactory receptor MOR261-6"  
/protein\_id="AAL61107.1"  
/db\_xref="GI:18480186"  
/translation="MGNNMTLITEFILGFPSPRMOMLFAFLSFLYFTLLNGTIL  
VGLICDSRLHTPMYFELSHLAIVDIAYACNTVPOMLVNLDVKKPISYAGCMOTFL  
FLTEFATECLLLVMSYDRYVAICHPLRYSAIMSWRVCSTMAVTSMIIGVLSLHLV  
LLPLPFCVSKVNHFECEITAILKLACADHLENTMVLGAVSVLVGFSSIVSYA  
CILGAILKIQSEEGQRKAFSTCSSHLGVGLPYGTAIMVYVGRHGSPEQKKYLLLF  
HSLFNPMLNPILYSLRNSDVKNLTKRVLRTORAL"  
BASE COUNT 199 a 249 c 200 g 285 t  
ORIGIN  
alignment\_scores:  
Quality: 1330.00 Length: 310  
Ratio: 4.634 Gaps: 0  
Percent Similarity: 92.581 Percent Identity: 81.935  
alignment\_block:  
US-09-898-586-24 x AY073444 ..  
Align seg 1/1 to: AY073444 from: 1 to: 933  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
1 ATGGGAAACAATATGACATTAATCACAGAGTTCATCTCTCGGATTTC 50  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34  
1 TCTCAGCCCGAGATGCAATGCTCTCTTGCCTCTTCTCTCTCTTCT 100  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
101 ATGCTTCACCTGCTGGGAAATGGAACCATCGTGGGCTCATCTGCC 150  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
151 GACTCCAGACTGCACACTCCATGTACTTCTCTGCTGTCTCACTGGCAT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84

:|||||  
201 CGTTGACATGGCTATGCTGCACACAGTGCCTGATGCTGAACC 250  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
251 TTTAGATCCAGTAGAGCCCATTTCCATGCTGTTGATGACACAGACT 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValme 117  
|||||  
301 TTTCTCTTTTGACATTTGCTATCAGAGAATGCCCTTCTCCTAGTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
351 GTCCCTATGACCGGATGTGCAATCTGTCAACCCACTCCGTTACTCTGCCA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
|||||  
401 TCATGAGCTGGAGAGTTTGCAGCACCACATGGCAGTGACTCTCGATTAT 450  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167  
|||||  
451 GGTGTCTGCTGCTTGTGATCCACCTGCTTACTTTACCTTACCTT 500  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
|||||  
501 CTGTGTGTCTCAGAAAGTGAACCACTTTCTGTGAATATTACAGCTATCC 550  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
551 TCAAACTTGCTGTGAGACACACACCTCAATGAGACGATGGTCTGGCT 600  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
|||||  
601 GGGGCAGTGTCTGTCTGTGGGCCATTCCTCAATGTGTCTCTTA 650  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
|||||  
651 TGCATGTATCCTTGTCTATCTGAGATCCAGTCAAGAGGGTCAAA 700  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
701 GGAAGCCCTTTCCACCTGCTCCACCTCTGTGTGATGTTGTT 750  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267  
|||||  
751 TATGGTACCGCATTTGCATGTACGTTGACCCAGACACGGAGGCCAAA 800  
267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnPrometL 284  
|||||  
801 GGAGCAGAAAGAAATATCTTTGCTGTTCAGAGTCTTTTAACCTATGC 850  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||||  
851 TCAACCCCTTGATCTATAGTCTCAGGAATCTGACGTGAAGAATACATTG 900  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
|||||  
901 AAAAGAGTGTGAGGACACAGAGACTTTG 930

seq\_name: gb\_pr:AC091768

seq\_documentation\_block:

LOCUS AC091768 104955 bp DNA linear PRI 09-JAN-2002

DEFINITION Homo sapiens BAC clone RP11-703N5 from 7, complete sequence.

ACCESSION AC091768

VERSION AC091768.4 GI:15145624

KEYWORDS HTG.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 104955)

AUTHORS Sulston,J.E. and Waterston,R.

TITLE Toward a complete human genome sequence  
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
MEDLINE 99063792  
REFERENCE 2 (bases 1 to 104955)  
AUTHORS Harris,A., Haakenson,W. and Spalding,L.  
TITLE The sequence of Homo sapiens BAC clone RP11-703N5  
JOURNAL Unpublished (2001)  
REFERENCE 3 (bases 1 to 104955)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 104955)  
AUTHORS Waterston,R.H.  
TITLE Direct Submission  
JOURNAL Submitted (09-AUG-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 5 (bases 1 to 104955)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT On Aug 9, 2001 this sequence version replaced gi:14550325.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu/gsc  
Contact: saplens@watson.wustl.edu  
----- Summary Statistics  
Center project name: H\_NH0703N05  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see  
http://www.nhgri.nih.gov/DIR/CTB/CHR7, send  
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanesse,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)

NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP4-669B10, 2000 bp overlap; the clone sequenced to the right is RP4-798C17, 2000 bp overlap. Actual start of this clone is at base position 123947 of RP4-669B10; actual end is at base position 64342 of RP11-798C17.



84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
29044 TCCTGCATCCAGCCCAACCCATCTCTTGGCTGGTCATGACCCAGATG 29093  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
29094 TTTCTGTTTGGAGTTTGGACATACAGAAATGTCCTCTCGTGGTGAT 29143  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
29144 GTCCATGATCGGTACGTGGCCATCTGCCACCCCTCCGATATTTACCA 29193  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150  
|||||  
29194 TCATGACCTGGAAGTCATGATCACTTTGGCATTGACTTCCGTGATTTTA 29243  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167  
|||||  
29244 GGAGTCTTATGGCCCTTGTCCATCTAGTGTACTGTACCACCTGCTCT 29293  
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
|||||  
29294 CTGTGCACCCCAAGAACTTAATCACTTTTCTGTGAATTAATGCTGTTC 29343  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
29344 TCAAACTTGGCTGTGGGATACCACCATTAATGAGTATAGTTTGGCA 29393  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerty 217  
|||||  
29394 GGGGAGTGTCTGTGGTGGGAGGCTCTTTTCCACTGTAATATCTTA 29443  
217 tMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlna 234  
||:::|  
29444 TGTTCAATATCTATGTGCATCTCTAAAGATCCAGTCAAGAGAGGGGTGCC 29493  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
:::|||||  
29494 AGAAAGCCTTCTCCATCTGCTCTCCACCTCTGTGTGGTGGACTCTTT 29543  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267  
|||||  
29544 TATGGCACAGCCATCATGATGTATGTGAGCCCCAGTATGAGAGCCCCAA 29593  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPrometL 284  
|||||  
29594 GGAGCAGAAGAATATCTCCCTGCTGTTCACAGCCTCTCAATCCCATGC 29643  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||||  
29644 TTAATCCCCCTAATTTAAGTCTTAGGAACAAGAGTCCAAGTACTCTTA 29693  
301 LysArgValLeuGlyValGluArg 308  
|||||  
29694 AAGAGGATGCTTGAAAGAAGAGA 29717

seq\_name: gb\_ro:AY073100

seq\_documentation\_block:  
LOCUS AY073100 933 bp DNA linear ROD 04-FEB-2002  
DEFINITION Mus musculus olfactory receptor MOR261-1 gene, complete cds.  
ACCESSION AY073100  
VERSION AY073100.1 GI:18479497  
KEYWORDS  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
1 (bases 1 to 933)  
AUTHORS Zhang, X. and Firestein, S.  
TITLE The olfactory receptor gene superfamily of the mouse  
JOURNAL Nat. Neurosci. 5 (2), 124-133 (2002)

PUBMED 11802173  
REFERENCE 2 (bases 1 to 933)  
AUTHORS Adams, M.  
TITLE Direct Submission  
JOURNAL Submitted (11-JAN-2002) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
FEATURES  
source location/Qualifiers  
1.933  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="6"  
1.>933  
/product="olfactory receptor MOR261-1"  
1.933  
/codon\_start=1  
/evidence=not\_experimental  
/product="olfactory receptor MOR261-1"  
/protein\_id="AA160763.1"  
/db\_xref="GI:18479498"  
/translation="MGNOTLITOFILLGFPLSPRMOLLFALFSLFYAFLLNGTI  
LGLICLDSRIHTPMYFFLSHLAIVDIAYACNTVPQMLVNLMDPAKPISFAGCMTOTFL  
FLVFANTECILLVMSYDRYVAICHPLRYTAISMNRVCVILVFTSWILGYLLALVHLV  
LILPLPCGSQKVNHFCEIILVLIACSDTRINELVLAGAVSVLVGPFSSIVSYA  
HILCALIKTSQGRQKAFSTCSHLGVGLFYGTAVMYIGPQHGKSNRQKYYLLLF  
HSLFNPMLNPILYSLRNKEYKSAIKRTLKEDTS"  
BASE COUNT 201 a 248 c 195 g 289 t  
ORIGIN  
alignment\_scores:  
Quality: 1246.00 Length: 304  
Ratio: 4.482 Gaps: 0  
Percent Similarity: 91.447 Percent Identity: 77.961  
alignment\_block:  
US-09-898-24 x AY073100 ..  
Align seg 1/1 to: AY073100 from: 1 to: 933  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
|||||  
1 ATGGGGGGAATCAGACCTTGATCAACAGAGTTCATCTCTGGGATTTC 50  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
||:::|  
51 TCTCAGCCCGAGAGATGACAGATGCTCTTGGCCCTCTTCCCTGTTCT 100  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
||:::|  
101 ATGCCCTGACCCCTGCTGGGAATGGGACCATCCCTGGGACTTATCTGCCCTG 150  
101 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
|||||  
151 GACTCCAGACTCCACACTCCCATGTACTTCTTCTGTCTCACCCTGGCCAT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
:|||||  
201 CGTTGACATTTGCCCTATGCCCTGCACAACAGTGGCCAGATGCTGTAACCC 250  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
||:::|  
251 TCATGATCAGCCCAAGCAATCTCTTTCGTGGTGCATGACACAGACC 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
301 TTTCTCTTTTGGAGTGTTCACACAGGAATGTCTCTACTGTTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
351 GTCCATGATCGATATGTAGCCATCTGCCACCCACATCCGATATACTGCTA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
|||||



```
401 TCATGAGCTGAGAGATTGTGTATCTCTGTCGACTTCCTGATTTTA 450
151 G1yValLeuSerLeuIleHisLeuValLeuLeuProLeuProph 167
|||||
451 GGAGTACTCTTGCCCTGCTCCTAGTATGTTGTTACCATTAACCTT 500
167 ecysArgProGlnIlySerHisPhePheCysGluIleLeuAlaVal 184
|||||
501 CTGCGGCTGCACAAAAGTGAATCACTTTTCTGTGAATTTATAGCTGCTC 550
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||||
551 TCAAACTGCTGTGTCAGATACCCGCATCAATGAGTTAATGTTTGCT 600
201 G1yAlaIleSerG1yLeuValG1yProLeuSerThrIleValValSerTy 217
|||||
601 GGGGCTGTGTCTGTGCTGTGCGACCATTTCTTCAATGTGATGCTCTTA 650
217 rmetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
|
651 TGCTCATATCTGTGTCCTGCTGCTGAGATTAAGTCAACAGCAGGCGGCC 700
234 rglysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
:|||||
701 AGAAAGCCCTCTCCACTGCTGCTCTCATCTCTGTTGTTGACCTTT 750
251 TyrG1yThrAlaIleIleMetTyrValG1yProArgTyrG1yAsnProLy 267
|||||
751 TATGTTACAGCCATTGTCATGTGACATTGGGCTTCAACACGGGAAGTCAA 800
267 sgluGlnIlySlyTyrLeuLeuPheHisSerLeuPheAsnPrometL 284
:|||||
801 TGAGCAGAGAAATATCTCTGCTATTCATAGCCTTTTAACTCTATGC 850
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
|||||
851 TCAACCCCTGATCTATAGTCTGAGACAAAGAGTCAAAAGTCTCTG 900
301 LysArgValLeu 304
|||||
901 AAGAGGACACTC 912
seq_name: gb_pr:AC074386
seq documentation block:
LOCUS AC074386 66610 bp DNA linear PRI 07-NOV-2001
DEFINITION Homo sapiens BAC clone RP11-464H1 from 7, complete sequence.
ACCESSION AC074386
VERSION AC074386.6 GI:13625507
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 66610)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 66610)
AUTHORS Cordum, H. and Cotton, M.
TITLE The sequence of Homo sapiens BAC clone RP11-464H1
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 66610)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 66610)
REFERENCE 4 (bases 1 to 66610)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (15-APR-2001) Genome Sequencing Center, Washington
```

```
REFERENCE University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 66610)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 66610)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Apr 15, 2001 this sequence version replaced gi:13431207.
COMMENT ----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wuston.wustl.edu
----- Summary Statistics
Center project name: H_NH0464H01
-----
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/CTB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenos, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)  
VECTOR: pBACe3.6  
NEIGHBORING SEQUENCE INFORMATION:  
The clone sequenced to the left is RP4-798C17, 200 bp overlap; the clone sequenced to the right is RP4-545C24, 200 bp overlap. Actual start of this clone is at base position 98638 of RP4-798C17; actual end is at base position 76962 of RP4-545C24.

There are polymorphic base pair differences in the overlap between the clone RP11-464H1 and RP4-545C24.

```
FEATURES
    source
        1..66610
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7"
            /clone="RP11-464H1"
            /clone_lib="RPCI-11"
            98..259
                /note="similar to ESP AI603950 (NID:94613112) ms16c11.y1"
```



```
misc_feature      98. .259 /note="similar to EST BE871418 (NID:g10320194)"
misc_feature      98. .259 /note="similar to EST BE383988 (NID:g9329353)"
misc_feature      98. .259 /note="similar to EST BE696685 (NID:g10083845)"
misc_feature      102. .259 /note="similar to EST BE383676 (NID:g9329139)"
misc_feature      114. .259 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      222. .259 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      297. .361 /note="similar to EST BG385901 (NID:g13279347)"
repeat_region     517. .646 /rpt_family="L2"
repeat_region     517. .646 /rpt_family="L1"
repeat_region     656. .951 /rpt_family="Alu"
repeat_region     1149. .1451 /rpt_family="Alu"
repeat_region     1484. .1806 /rpt_family="L2"
misc_feature      1743. .2035 /note="similar to EST AA078615 (NID:g1838096)"
repeat_region     1942. .2074 /rpt_family="7SK"
repeat_region     2159. .2234 /rpt_family="MIR"
repeat_region     2701. .2896 /rpt_family="MIR"
repeat_region     2908. .2949 /rpt_family="MIR"
repeat_region     3044. .3349 /rpt_family="Alu"
misc_feature      4064. .4319 /note="similar to EST BF027677 (NID:g10735389)"
misc_feature      4065. .4272 /note="similar to EST D77287 (NID:g1597007)"
misc_feature      4066. .4319 /note="similar to EST BG385901 (NID:g13279347)"
misc_feature      4066. .4319 /note="similar to EST BE383676 (NID:g9329139)"
misc_feature      4066. .4319 /note="similar to EST BE383988 (NID:g9329353)"
misc_feature      4066. .4265 /note="similar to EST AI603950 (NID:g4613112) ms16c11.y1"
misc_feature      4066. .4127 /note="similar to EST BE696685 (NID:g10083845)"
misc_feature      4066. .4319 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      4066. .4319 /note="similar to EST BE871418 (NID:g10320194)"
misc_feature      4164. .4319 /note="similar to EST AI910192 (NID:g6500872)"
misc_feature      4511. .4575 /note="similar to EST BE383988 (NID:g9329353)"
misc_feature      4511. .4575 /note="similar to EST AI910192 (NID:g6500872)"
misc_feature      4511. .4575 /note="similar to EST BG385901 (NID:g13279347)"
misc_feature      4511. .4575 /note="similar to EST BE383676 (NID:g9329139)"
misc_feature      4511. .4575 /note="similar to EST BF027677 (NID:g10735389)"
misc_feature      4511. .4575 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      4511. .4575 /note="similar to EST BE871418 (NID:g10320194)"
misc_feature      5009. .5066 /note="similar to EST BE383988 (NID:g9329353)"
misc_feature      5009. .5066 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      5009. .5066
```

```
misc_feature      /note="similar to EST BE383676 (NID:g9329139)"
misc_feature      5009. .5066 /note="similar to EST AI910192 (NID:g6500872)"
misc_feature      5009. .5066 /note="similar to EST BG385901 (NID:g13279347)"
misc_feature      5009. .5066 /note="similar to EST BE871418 (NID:g10320194)"
misc_feature      5009. .5066 /note="similar to EST BF027677 (NID:g10735389)"
misc_feature      5598. .5674 /note="similar to EST BF027677 (NID:g10735389)"
misc_feature      5598. .5670 /note="similar to EST BE383676 (NID:g9329139)"
misc_feature      5598. .5729 /note="similar to EST AI910192 (NID:g6500872)"
misc_feature      5598. .5834 /note="similar to EST BG385901 (NID:g13279347)"
misc_feature      5598. .5670 /note="similar to EST BE410379 (NID:g9346829)"
misc_feature      5696. .6059 /note="similar to EST BE383988 (NID:g9329353)"
misc_feature      5718. .6272 /note="similar to EST BF328558 (NID:g11299410)"
misc_feature      5800. .5923 /note="similar to EST AW784785 (NID:g7841561)"
```

alignment\_scores:      Quality: 1138.00      length: 304  
                         Ratio: 4.246      Gaps: 0  
Percent Similarity: 88.158      Percent Identity: 72.039

alignment\_block:  
US-09-898-586-24 x AC074386 ..

Align seg 1/1 to: AC074386 from: 1 to: 66610

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
21158 ATGCTGAAATAATCAGACATGCTCAGAGATTCCTCTACTGGATTCT 21207
17 oValGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheT 34
21208 CCGGCCCAAGGATTCAGATGCTCTTTGGGCTCTTCTCCCTGTTCT 21257
34 yrrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
21258 ATGCTTCACCCCTGCTGGGAATGGACCATCGGGCTCATCTGACTG 21307
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
21308 GACTCCAGACTCCACACCCCATGTACTTCTCTCTCTCACACCTGGCCGT 21357
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
21358 CGTCACATCGCCCTATGCTGCACACACAGTGCCTGATGCTGGTAACC 21407
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
21408 TCCTGCATCCAGCCCAAGCCCATCTCCTTGTGCTGCATGACATAGACC 21457
101 pheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117
21458 TTTCTCTTTTGTAGTTTGGACATACTGAATGCTCTGTGTGCTGAT 21507
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
21508 GTCCCTACGATCGGTACGTGGCCATCTGCCACCCCTCTCCGATATTTCATCA 21557
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTyrPheThr 150
21558 TCATGACCTGGAAGTCTGCATCACTCTGGCCATCACTTCTGACATGT 21607
```

```

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167
    ||| |||:::||||: |||::||| ||||||
21608 GGCTCCCTCCCTGGCTATGCTCATGTGAGCCCTCATCTAAGACTGCCCT 21657
    ||| |||:::|||| |||||| ||| |||:::||||
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184
    ||| |||:::|||| ||| ||| ||| ||| ||| ||| ||| |||
21658 TTGTGGGCTCGTGTGAATCAACCACTTCTCTGTGAATCCTGTCTGCC 21707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 eulysleuAlaCysAlaAspThrHisIleasnGluAsnmetValLeuAla 200
    ||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21708 TCAGGCTGGCTGTGTGATACCTGGCTCAACAGGTGGTGCATCTTTGCA 21757
    ||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSery 217
    ::::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21758 GCGTGCAATGTTCAATCCGTGGGACCACTTGCCCTGGTGTCTCTCCTA 21807
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21808 CTCACACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGAGGCCGCA 21857
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 rGlyAlaPheCysThrCysPheSerHisIleuCysValIleGlyLeuPhe 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
21858 GAAAGGCTTCTCCACCTGCTCCCTCCACCTGCGGTAGTGGGACTCTTC 21907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
    ::|||:::||||: ||| ||| ||| ||| ||| ||| ||| ||| |||
21908 TTTGGCAGCGCATGTCATGTACATGCGCCCTAAGTCCCGCATCCCTGA 21957
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMet 284
    :||| |||:::|||| ||| ||| ||| ||| ||| ||| ||| |||
21958 GGAGCAGCAGAAAGTCTCTTTCTATTTTACAGTTCPTTCAACCCGATGC 22007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
22008 TAAACCCCTGATTTACAACTGAGGAATGTAGAGGTCAAGGGTGCCCTG 22057
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
301 LysArgValLeu 304
    ::|||:::||||
22058 AGGAGAGCACTG 22069

```

OM of: US-09-898-586-24 to: EST:\* out\_format : pfs  
Date: Aug 21, 2002 7:47 PM

About: Results were produced by the Gencore software, version 4.5,  
Copyright (c) 1993-2000 Compugen Ltd.

## Command line parameters:

-MODE=frame+\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US09898586/runat\_16082002\_205735\_20538/app-query.fasta\_1.373  
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOPCL=0.000 -LOPEXT=0.000 -OGAPOP=4.500  
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEXT=7.000 -START=1 -MATRIX=bls62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09898586@cgn1\_1\_2668  
-NCPU=6 -ICPU=3 -LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30  
-NO\_XLPXY -WAIT -THREADS=1

## Search information block:

Query: US-09-898-586-24  
Query length: 310  
Database: EST:\*  
Database sequences: 13736207  
Database length: -1841457050  
Search time (sec): 1599.010000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
gb_htc:BC016940	+ 1138.00	1904.12	8.2e-97	2021	BC016940 Homo sapiens, Similar
gb_est1:AI148854	+ 794.00	1333.22	5.2e-65	619	AI148854 qc70a02.x1 Soares, plia
gb_htc:AF327904	+ 792.00	1328.62	9.3e-65	705	AF327904 Homo sapiens FKSG35 (H
gb_est2:BE856570	+ 792.00	1328.03	1.0e-64	751	BE856570 7f64e08.x1 Soares_NSF
gb_est1:AI339565	- 790.00	1329.32	8.5e-65	456	AI339565 qk67e07.x1 NCI CGAP_C
gb_est2:BE543056	- 762.00	1279.69	4.9e-62	580	BE543056 601068985F1 NIH_MGC_12
gb_est1:AI339573	- 754.00	1268.90	2.0e-61	433	AI339573 qk67f03.x1 NCI CGAP_C
gb_est2:BG984967	- 752.00	1265.30	3.1e-61	443	BG984967 IL5-CN0068-100401-434-
gb_est1:AI079550	+ 750.00	1257.13	8.9e-61	738	AI079550 oz04f08.x1 Soares_feta
gb_gss:BH111304	+ 742.50	1243.83	4.9e-60	788	BH111304 RPCI-24-367N6.TJ RPCI-
gb_est2:BF974217	- 687.00	1151.44	6.9e-55	671	BF974217 602243851F1 NIH_MGC_48
gb_est2:BF116115	+ 678.00	1135.39	5.4e-54	732	BF116115 7n77g04.x1 NCI CGAP_OV
gb_est1:AI340119	- 643.00	1082.66	4.7e-51	367	AI340119 qk64f08.x1 NCI CGAP_C
gb_est1:AI285267	- 602.00	1013.45	3.3e-47	361	AI285267 q138c08.x1 NCI CGAP_C
gb_est2:H62441	+ 593.00	996.67	2.9e-46	426	H62441 yr79e08.r1 Soares fetal
gb_est2:H62445	+ 584.00	981.16	2.1e-45	439	H62445 yr79e08.r1 Soares fetal
gb_est1:BE179276	+ 581.00	971.84	6.9e-45	690	BE179276 RCI-HT0615-290300-021-
gb_htc:AK017036	+ 580.00	965.24	1.6e-44	1394	AK017036 Mus musculus adult ma
gb_htc:AK016560	+ 580.00	956.16	5.2e-44	3063	AK016560 Mus musculus adult ma
gb_est1:AW299289	+ 560.00	939.35	4.5e-43	499	AW299289 xs38f05.x1 NCI CGAP_K1
gb_est1:BE174029	- 554.00	879.55	1.4e-42	432	BE174029 QV1-HT0572-280200-089-
gb_est1:BE182734	- 522.50	879.73	9.4e-40	332	BE182734 RC3-HT0650-130400-011-
gb_gss:AO600752	+ 517.50	867.06	4.8e-39	520	AO600752 HS_5346_B2_F04.T7 RPCI
gb_est2:BI020419	+ 509.00	854.22	2.5e-38	441	BI020419 CM3-MT0292-120101-616-
gb_est2:BG193339	+ 498.00	830.48	5.2e-37	762	BG193339 RST12467 Athersys RAGB
gb_est2:BG193339	+ 489.50	815.04	3.8e-36	853	BG193339 CH230-46N15.TV CHORI-2
gb_gss:BH331857	- 474.00	794.84	5.0e-35	449	BH331857 ypolh11.r1 Soares breast
gb_est2:H39851	- 466.50	776.78	5.1e-34	796	H39851 ypolh11.r1 Soares breast
gb_est2:BG197640	+ 466.00	769.98	1.2e-33	1501	AK016338 Mus musculus adult ma
gb_htc:AK016338	+ 464.00	773.41	7.8e-34	726	AZ255734 RPCI-23-16619.TJ RPCI-
gb_gss:AZ255734	- 455.50	758.16	5.5e-33	797	AZ255734 RPCI-24-222G18.TJ RPCI-
gb_gss:AZ2909618	+ 455.00	759.34	4.8e-33	642	AZ2909618 2M0241J24R Mouse 10kb
gb_gss:AZ2969227	- 446.00	743.91	3.4e-32	656	BH078380 RPCI-24-349E9.TV RPCI-
gb_gss:BH078380	- 446.00	742.78	4.0e-32	740	AZ607393 LM0429N1SR Mouse 10kb
gb_gss:AZ607393	- 442.50	737.49	7.8e-32	692	AZ607393 RPCI-23-27B20.TV RPCI-
gb_gss:AZ2090606	- 442.50	735.65	9.9e-32	842	AZ275747 RPCI-24-98112.TVB RPCI-
gb_gss:AZ275747	- 442.00	737.81	7.5e-32	611	AZ103967 RPCI-23-33G18.TV RPCI-
gb_gss:AZ103967	- 438.50	730.94	1.8e-31	534	AZ64867 LM0545E05R Mouse 10kb
gb_gss:AZ353949	+ 437.00	730.61	1.9e-31	679	AG167722 Pan troglodytes DNA, d
gb_gss:AG167722	+ 434.00	723.29	4.8e-31	679	

gb\_gss:AZ913406 - 432.50 719.28 8.1e-31 794 1 AZ913406 RPCI-24-165C12.TV R  
gb\_gss:AZ019257 - 432.00 720.62 6.8e-31 629 1 AZ019257 RPCI-23-266H21.TV R  
gb\_gss:BH279965 + 427.00 710.79 2.4e-30 728 1 BH279965 CH230-38K1.TJ CHORI  
gb\_gss:BH359995 + 427.00 710.43 2.5e-30 756 1 BH359995 CH230-126I6.TJ CHOR  
gb\_gss:AZ765752 + 425.00 708.73 3.1e-30 632 1 AZ765752 LM0562E19R Mouse 10

seq\_name: gb\_htc:BC016940

seq\_documentation\_block:

LOCUS BC016940 2021 bp mRNA linear HTC 09-NOV-2001  
DEFINITION Homo sapiens, similar to olfactory receptor, family 2, subfamily A,  
member 4, clone IMAGE:4424116, mRNA.

ACCESSION BC016940  
VERSION BC016940.1 GI:16877381

KEYWORDS HTC.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2021)  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
COMMENT Contact: MGC help desk  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Sequencing Group at the Stanford Human Genome  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcd@paxil.stanford.edu](mailto:mcd@paxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,  
R. M.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAC Plate: 27 Row: 1 Column: 22  
This clone was selected for full length sequencing because it  
passed the following selection criteria: Hexamer frequency ORF  
analysis, Genomescan gene prediction, similarity but not identity  
to protein  
This clone has the following problem: frame shifted.

FEATURES  
source location/Qualifiers  
1..2021  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4424116"  
/tissue\_type="Kidney, hypernephroma"  
/clone\_1lb="NIH\_MGC\_89"  
/lab\_host="DH10B"  
/note="Vector: pCMV-Sport6"

BASE COUNT 470 a 546 c 427 g 578 t  
ORIGIN

alignment\_scores:  
Quality: 1138.00 Length: 304  
Ratio: 4.246 Gaps: 0  
Percent Similarity: 88.158 Percent Identity: 72.039

alignment\_block:  
US-09-898-586-24 x BC016940 ..

Align seg 1/1 to: BC016940 from: 1 to: 2021

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
|||||  
429 ATGGTGAATAATCAGACAAATGGTCAAGAGTTCCTCTACTGGAATTTC 478

17	oValGIyPrOArGIleGIleMetLeuLeuPheGlyLeuPheSerLeuPheT	34
479	CTGGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCTCCCTGTTCT	528
34	yValAlpheThrLeuLeuGLyAsnGlyThrIleLeuGLyLeuIleSerLeu	50
529	ATGCTTTCACCCCTGCTGGGAATGGACCATCTCTGGGGCTCATCTCACTG	578
51	AspSerArGLeuHisAlaProMetTyTPhePheLeuSerHisLeuAlaVa	67
579	GACTCCAGACTCCACACACCCCATGTACTTCTTCTCTCACACCTGGCCGT	628
67	lValAspIleAlaTyAlaIaCysAsnThrValProArgMetLeuValAsnL	84
629	CGTCACATCGCCCTATGCCTGCACACAGTGGCCCAAGATGCTGTGAACC	678
84	euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr	100
679	TCCTGCATCCAGCCAAAGCCCATCTCTTGTGCGCATGACATAGACC	728
101	PhelLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValAlMe	117
729	TTTCTCTTTTGAGTTTGGACATACATGAATGCCTCTCTTGTGTGAT	778
117	tSerTyArSpleuTyValAlaIaIeCysHisProLeuArgTyTrLeuAlaI	134
779	GTCCCTACGATCGGTACGTGGCCATCTGGCCACCTCTCCGATATTTCATCA	828
134	leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr	150
829	TCATGACCTGGAAAGTCTGCATCACTCTGGCCATCACTCTCGACATGT	878
151	GlyValLeuLeuSerIleuIleHisLeuValLeuLeuLeuProLeuProPh	167
879	GGCTCCCTCCTGGCTATGTCATGTGAGCCCTCATCTTAAGACTGGCCCT	928
167	ecysArgProGlnLysIleTyHisPhePheCysGluIleLeuAlaValL	184
929	TTGTGGGCGCTGTAATCAACCACTTCTCTGTGAATCCGTCTGTCTCC	978
184	eulLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla	200
979	TCAGGCTGGCCTGTGCTGATACCTGGCTCAACCAAGGTGCTCATCTTTGCA	1028
201	GIyAlaIleSerGIyLeuValGIyProLeuSerThrIleValAlaSerTy	217
1029	GCCTGCATGTTCATCCTGTGTGGGACCACTCTGCTGTGCTGTCTTCTTA	1078
217	rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArGluValGlnA	234
1079	CTCACACATCTCTGGCGGCCATCTTGAGGATCCAGTCTGGGGAGGGCCGCA	1128
234	rGlyAlaIaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe	250
1129	GAAAGCCCTTCTCCACCTGTCTCCCACTCTGCGTAGTGGGACTCTTC	1178
251	tyrGIyThrAlaIleIleMetTyValGIyProArgTyrGIyAsnProLy	267
1179	TTTGGCAGCGGCATCGTCATGTACATGAGCCCTAAGTCCCGCCATCTGA	1228
267	sGluGlnLysLysTyTrLeuLeuPheHisSerLeuPheAsnProMetL	284
1229	GGAGCAGCAGAAAGTCTTTTCTTAATTTCACAGTCTCTTCAACCCGATGC	1278
284	euaSnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu	300
1279	TAAACCCCTGATTTTACAAACCTGAGGATGTAGAGGTCAAGGGTGCCTTG	1328
301	LysArgValLeu	304
1329	AGGAGGCACTG	1340

```
seq_name: gb_estl:A1148854

seq_documentation_block:
LOCUS A1148854 619 bp mRNA linear EST 28-OCT-1998
DEFINITION qc70a02.x1 Soares_placenta_8to9weeks_2NDHP8to9W Homo sapiens cDNA clone IMAGE:1714922 3' similar to SW:OLF7_MOUSE P34984 OLFATORY RECEPTOR-LIKE PROTEIN K7 ;, mRNA sequence.
ACCESSION A1148854
VERSION A1148854.1 GI:3677323
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 619)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgaps-r@mail.nih.gov
This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1098 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 432.
FEATURES
    source
        location/Qualifiers
            1..619
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1714922"
                /clone_lib="Soares_placenta_8to9weeks_2NDHP8to9W"
                /dev_stage="two placentae: one from 8 weeks and another from 9 weeks post conception"
                /lab_host="DH10B (ampicillin resistant)"
                /note="Organ: placenta; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCATCTGAAGTCGGACGCCGCCGCGATTCTTTTCTTTTCTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 112 a 191 c 131 g 183 t 2 others
ORIGIN
alignment_scores:
    Quality: 794.00      Length: 198
    Ratio: 4.511         Gaps: 0
Percent Similarity: 88.889 Percent Identity: 76.768

alignment_block:
US-09-898-586-24 x A1148854 ..

Align seg 1/1 to: A1148854 from: 1 to: 619
1 MetcGLyAsPasnleThrSerleThrgluPhelLeuenglyPhePr 17
   |||          |||          :::|||||
26 ATGCTGAANAATCAGACAAGCGTCACAGAGTTCCTCTACTGGGATTTCCT 75
17 oValgLyProArgileglmetLeuLeupheglyLeuPheSerLeuPheT 34
   :::|||||
76 CCTGGGCCCAAGGATTCAAGATGCCTCTTGGGCTCTTCCCTGTCTCT 125
34 yValpheThrLeuLeuglyAsnclYThrIlleLeuglyLeuSerLeu 50
   |||
126 AGTCTTCACCCCTGCTGGGGAATGGAGCACCATCTGGGGCTCATCTCACTG 175
51 AsperArgleuHisAlaPrometTyrrPhepheLeuSerHisLeuAlaVa 67
   |||
176 GACTCAGACTCCACACCCCCCATGTACTTCTTCTCTCACACCTGGCCGT 225
```

```
67 lvalaspillealatyralacysasnthrvalproargmetleuvalasnL 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
226 CGTCAACGTCGGCTATGCTGCACACACAGTGCCTGCTGTGAACC 275
84 euleuHisProAlalysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
276 TCCTGCATCCAGCCCAAGCCCATCTCCCTTGGCTGGCTGCATGACATAGACC 325
101 pheleupheserThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
326 TTTCTCTTTTGTAGTTTTCACATACTGAATGCCCTCCTGTTGGTCTGAT 375
117 tSerTyraSPleuTyraValAlaIleCysSHisProLeuArgTyrlleuAlaI 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
376 GTCTACGATCGGTACCTGGCCATCTGCCACCCCTCTCCGATATTTCATCA 425
134 lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
426 TCATGACCTGMAAAGTCTGCATCACTCTGGCCATCACTTCCGTGACATGT 475
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
476 GGCTCCCTCCTCGGCTATGNCATGTGAGCCTCATCTTAAGACTGCCCTT 525
167 ecysArgProGlnLysIleTyrlHisPhePheCysGluIleLeuAlaValL 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
526 TTGTGGGCGCTCGTGAATCAACCACTTCTTGTGAATCCGTGTGTGTC 575
184 eulysleuAlaCysAlaSPThrHisIleasnGluasmetVal 198
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
576 TCAGGCTGGGCTGTGCTGATACCTGGCTCAACCAAGTGTGTCATC 619

seq_name: gb_hlc:AF327904

seq_documentation_block:
LOCUS AF327904 705 bp mRNA linear HTC 13-JUL-2001
DEFINITION Homo sapiens FKSG35 (FKSG35) mRNA, complete cds.
ACCESSION AF327904
VERSION AF327904.1 GI:12276181
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 705)
Wang, Y.-G. and Gong, L.
Cloning of FKSG35, a novel gene located on human chromosome 7
JOURNAL
REFERENCE
2 (bases 1 to 705)
Wang, Y.-G.
Direct Submission
Submitted (13-DEC-2000) Beijing Fengkesheng Function Gene
Technology Ltd., 4 Tou Tiao Lu Chang Street, Xuanwu District,
Beijing 100050, P.R. China
FEATURES
source
1..705
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q33-q35"
/tissue_type="liver"
1..705
/gene="FKSG35"
133..429
/gene="FKSG35"
/codon_start=1
/product="FKSG35"
/protein_id="AAG50283.1"
/db_xref="GI:12276182"
/translation="MVKNQTVTEFLLGFLGPRIGMLFGLFSFYVFTLLNGTI
LGLISLDSRLHPMYFLSHLAVVNIAACNTVPQMLVNLHPAKPISFAGCMT"
BASE COUNT 156 a 201 c 132 g 216 t
```

```
ORIGIN
alignment_scores:
Quality: 792.00 Length: 191
Ratio: 4.632 Gaps: 0
Percent Similarity: 89.529 Percent Identity: 79.581
alignment_block:
US-09-898-586-24 x AF327904 ..
Align seg 1/1 to: AF327904 from: 1 to: 705

1 MetGlyaspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
||| ||| ||| ::||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 ATGCTGAATAATCAGACAAATGTCACAGAGTTCCTCTACTGGGATTTC 182
17 ovalGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
183 CCTGGGCCCAAGGATTCAAGATGCTCCTCTTGGGCTCTTCTCCCTGTCT 232
34 yrvAlPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
233 ATGCTTCACCCCTGCTGGGGAATGGGACCATCTGGGGCTCATCTCACTG 282
51 AspSerArgLeuHisAlaProMetTyrlPhePheLeuSerHisLeuAlaVa 67
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
283 GACTCCAGACTCCACACCCCATGTACTTCTTCTCTCACACCTGGCCGT 332
67 lvalaspillealatyralacysasnthrvalproargmetleuvalasnL 84
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
333 CGTCAACATCGCCTATGCTGCAACACAGTGCCTGAGATGCTGTGAACC 382
84 euleuHisProAlalysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
383 TCCTGCATCCAGCCCAAGCCCATCTCTTGTGCTGGCTGCATGACATAGACC 432
101 pheleupheserThrPheAlaValThrGluCysLeuLeuLeuValValMe 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
433 TTTCTCTTTTGTAGTTTTCACATACTGAATGCCCTCCTGTGTGCTGAT 482
117 tSerTyraSPleuTyraValAlaIleCysSHisProLeuArgTyrlleuAlaI 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
483 GTCTACGATCGGTACGTGGCCATCTGCCACCCCTCTCCGATATTTCATCA 532
134 lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
533 TCATGACCTGGAAAGTCTGCATCACTCTTGGCATCACTTCTTGACACATGT 582
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
583 GGCTCCCTCCTGGCTATGTCATGTGAGCCCTCATCTTAAGACTGCCCTT 632
167 ecysArgProGlnLysIleTyrlHisPhePheCysGluIleLeuAlaValL 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
633 TTGTGGGCGCTCGTGAATCAACCACTTCTTGTGAATAAACCCTGTGTGTC 682
184 eulysleuAlaCysAlaAspThr 191
::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
683 CTCAGCTGGCCTGTGCTGATACC 705

seq_name: gb_est2:BE856570

seq_documentation_block:
LOCUS BE856570 751 bp mRNA linear EST 29-SEP-2000
DEFINITION 7f64e08.x1 Soares NSF_F8_9W_OT_PA_P-S1 Homo sapiens cDNA clone
IMAGE:3299462 3' similar to TR:095047 095047 WUGSC:H_PJ0988615.2
PROTEIN: ;, mRNA sequence.
ACCESSION BE856570
VERSION BE856570.1 GI:10369724
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
```





## ORIGIN

## alignment\_scores:

Quality: 790.00 Length: 151  
Ratio: 5.232 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-898-586-24 x AI339565/rev ..

Align seg 1/1 to reverse of: AI339565 from: 1 to: 456

```
31 SerLeuPheTyrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLe 47
|||||
454 TCCCTGTTCTACGCTTTCACCCCTGCTGGGGAACGGACCATACTGGGCT 405

47 uIleSerLeuAspSerArgLeuHisAlaProMetTyrPhePheLeuSerH 64
|||||
404 CATCTCACTGGACTCCAGACTGCACGCCCCCATGTACTTCTCTCTCAC 355

64 lslLeuAlaValValAspIleAlaTyrAlaCysAsnThrValProArgMet 80
|||||
354 ACCGTGGCGGTCGTGCACATCGCCTACGCGCTGCACACAGGTGCCCGGATG 305

81 LeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGlyArgMe 97
|||||
304 CTGGTGAACCTCCTGCATCCAGCCAGCCCACTCTTGGCGGCGGCAT 255

97 tMetGlnThrPheLeuPheSerThrPheAlaValThrGluCysLeuLeuL 114
|||||
254 GATGCAGACCTTCTGTTTCCACTTTTGGCTGCACAGAATGTCCTCTCC 205

114 euValValMetSerTyrAspLeuTyrValAlaIleCysHisProLeuArg 130
|||||
204 TGGTGTGATGTCCTATGATCTGTACGTGGCCACTGCGACCCCTCCGA 155

131 TyrLeuAlaIleMetThrTrpArgValCysIleThrLeuAlaValPheSe 147
|||||
154 TATTTGGCCATCATGACCTCGAGAGAGTCTGCATCACCCCTCGGTACTTC 105

147 rTyrThrThrGlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuP 164
|||||
104 CTGACCACTGAGAGTCTTATTCCTTGATTCACACTTGTACTTCTTAC 55

164 roLeuProPheCysArgProGlnLysIleTyrHisPhePheCysGluIle 180
|||||
54 CTTTACCCTTCTGTAGGCCCCAGAAAATTTATCACCTTTTTTGTGAATC 5

181 Leu 181
|||
4 TTA 2
```

seq\_name: gb\_est2:BE543056

## seq\_documentation\_block:

LOCUS BE543056 580 bp mRNA linear EST 09-AUG-2000  
DEFINITION 601068985F1 NIH\_MGC\_12 Homo sapiens cDNA clone IMAGE:3455324 5',  
mRNA sequence.

## ACCESSION

BE543056  
BE543056 GI:9771701

## KEYWORDS

## SOURCE

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 580)  
NIH-MGC <http://mgc.nci.nih.gov/>.

## AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.

Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM8441 row: i column: 21

High quality sequence start: 3

High quality sequence stop: 580.

## FEATURES

## source

1..580

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:3455324"

/clone\_lib="NIH\_MGC\_12"

/tissue\_type="cervical carcinoma cell line"

/lab\_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.4 kb. Library prepared by Life  
Technologies."

BASE COUNT 96 a 189 c 125 g 170 t  
ORIGIN

## alignment\_scores:

Quality: 762.00 Length: 193  
Ratio: 4.379 Gaps: 0  
Percent Similarity: 90.155 Percent Identity: 77.202

## alignment\_block:

US-09-898-586-24 x BE543056 ..

Align seg 1/1 to: BE543056 from: 1 to: 580

```
22 lIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValPheThrLe 38
|||||
3 ATTCAAGATGCTCCTTGGGCTCTTCCCTGCTGCTATGCTTACACCT 52

38 uLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSerArgLeuH 55
|||||
53 GCTGGGAATGGGACCATCTGGGGCTCATCTCAGTGGACTCCAGACTCC 102

55 lAlaProMetTyrPhePheLeuSerHisLeuAlaValAlaAspIleAla 71
|||||
103 ACACCCCATGTACTCTCTCCCTCACACCTGGCGCGTCCACACATCGCC 152

72 TyrAlaCysAnThrValProArgMetLeuValAsnLeuLeuHisProAl 88
|||||
153 TATGCCCTGCAACACAGTGGCCCGAGATGCTGTGAACCTCTGCATCCAGC 202

88 aLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeuPheSerT 105
|||||
203 CAAGCCCATCTCCTTGTGGCTGCATGACATAGACCTTCTCTTTTGA 252

105 hrPheAlaValThrGluCysLeuLeuValValMetSerTyrAspLeu 121
|||||
253 GTTTTGCACATACTGAATGCCCTCTGTTGGTGTGATGCTTACGATCGG 302

122 TyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMetThrTrpAr 138
|||||
303 TACGTGGCCATCTGCGCACCCCTCCGATATTTCATCATCATGACCTGGAA 352

138 gValCysIleThrLeuAlaValThrSerTyrThrGlyValLeuLeuS 155
|||||
353 AGTCTGCATCACTCTGGCCATCATCTCTGGACATGTGCTCCCTCTGG 402

155 erLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgProGln 171
|||||
403 CTATGCTCATGTGAGCCTCATCTTAAGACTGCCCTTTTGTGGGCTCGT 452

172 LysIleTyrHisPhePheCysGluIleLeuAlaValLeuLysLeuAlaCy 188
|||||
```

```
453 GAAATCAACCACTCTCTCTGTGAAATCCTGTCTGTCTCAGGCTGGCCTG 502
188 salaasPThrHisIleasnGLuasMetValleuAlaGlyAlaIleSerg 205
|||||:::|||||
503 TGTGATACCTGGCTCAACACAGGTGTCATCTTTGCCGCTGCATGTTCAT 551
205 lyleuValGlyProleuSerThrIleVal 214
|||||:::
552 C.CTGTGGGACCACTCTGCCCTGGTGTG 580

seq_name: gb_est1:AI339573

seq_documentation_block:
LOCUS AI339573 433 bp mRNA linear EST 13-FEB-1999
DEFINITION qk67f03.x1 NCI_CGAP_Co8 Homo sapiens CDNA clone IMAGE:1874045 3'
similar to SW:OLF7_MOUSE P34984 OLFATORY RECEPTOR-LIKE PROTEIN K7
; mRNA sequence.
ACCESSION AI339573 GI:4076500
VERSION AI339573
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 433)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 677 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 411.
FEATURES
Source
1.433
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1874045"
/clone_lib="NCI_CGAP_Co8"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo."
BASE COUNT 127 a 91 c 139 g 76 t
ORIGIN

alignment_scores:
Quality: 754.00 Length: 144
Ratio: 5.236 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
us-09-898-586-24 x AI339573/rev ..
Align seg 1/1 to reverse of: AI339573 from: 1 to: 433

38 LeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSerArgIle 54
```

```
|||||
433 CTGCTGGGGAACGGACCACTACTGGGGCTCATCTCTCACTGACCTCCAGACT 384
54 uHisAlaProMetTyrPhePheLeuSerHisLeuAlaValAlaIleA 71
|||||
383 GCACGCCCCATGTACTCTCTCTCTCACACCTGGCGGTGACATGCG 334
71 laTyrAlaCysAsnThrValProArgMetLeuValAsnLeuHisPro 87
|||||
333 CCTACCGCTGCACACACCGGTGCCCGGATGCTGTGAACCTCTCATCCA 284
88 AlAlaYsProIleSerPheAlaGlyArgMetMetGlnThrPheLeuPhe 104
|||||
283 GCCAAGCCCATCTCTTCGGGGCCGCATGATGCAGACCTTCTGTTTTC 234
104 rThrPheAlaValThrGluCysLeuLeuValValMetSerTyrAspL 121
|||||
233 CACTTTCTGTCTCACAGAAAGTCTCTCTCGGTGGTGATGTCCTATGATC 184
121 euTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMetThrTrp 137
|||||
183 TGPACGTGGCCATCTGCCACCCCTCCGATATTTGGCCATCATGACCTGG 134
138 ArgValCysIleThrLeuAlaValThrSerTrpThrGlyValIleule 154
|||||
133 AGAGTCTGCATCACCCCTCGGGGTGACTTCTTGACACCACTGAGATCTTTT 84
154 uSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgProg 171
|||||
83 ATCCTGTGATTCATCTGTGTACTTCTTACCTTACCTTCTGTAGGCCCC 34
171 lnlYsIleTyrHisPhePheCysGluIleLeu 181
|||||
33 AGAAATTTATCACTTTTGTGAAATCTTA 2

seq_name: gb_est2:BG984967

seq_documentation_block:
LOCUS BG984967 443 bp mRNA linear EST 12-JUN-2001
DEFINITION IL5-CN0068-100401-434-d09 CN0068 Homo sapiens CDNA, mRNA sequence.
ACCESSION BG984967
VERSION BG984967.1 GI:14387702
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 443)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagal,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=IL5&t2=IL5-CN0068-
100401-434-d09&t3=2001-04-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 363.
FEATURES
Location/Qualifiers
```





```
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117
|||||
484 TTTCTCTTTTGTGAGTTTGTGACATACATGAAATGCCTCTGTTGGTGTGAT 533
|||||
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
|||||
534 GTCCCTACGATCGGTACGTGGCCATCTGCCACCCCTCTCCGATATTTCATCA 583
|||||
134 leMetThrTrpArgValAlaCysIleThrLeuAlaValThrSerTrpThr 150
|||||
584 TCATGACCTGGAAAGTCTGCATCAGCTGGCCATCAGCTCTGGACATGT 633
|||||
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167
|||||
634 GGCTCCCTCCCTGGCTATGTGTCATGTGAGCCCTCATCTAGACTGGCCCT 683
|||||
167 ecysArgProGlnIleTyrHisPhePheCysGluIleLeuAlaVal 184
|||||
684 TTTGTGGCCCTGTGAAATCAACACACTTCTGTGAAATCCTTGCTGTCC 733
|||||
184 eulys 185
734 TCAG 738
seq_name: gb_gss:BH111304
seq_documentation_block:
LOCUS BH111304 788 bp DNA linear GSS 19-JUL-2001
DEFINITION RPCI-24-367N6.TJ RPCI-24 Mus musculus genomic clone RPCI-24-367N6,
DNA sequence.
ACCESSION BH111304
VERSION BH111304.1 GI:14946319
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 788)
AUTHORS Zhao,S., Niemman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE Mouse BAC End Sequences from Library RPCI-24
JOURNAL Unpublished (1995)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..788
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-367N6"
/clone_lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT 168 a 217 c 162 g 241 t
```

```
ORIGIN
alignment_scores:
Quality: 742.50 Length: 264
Ratio: 3.486 Gaps: 5
Percent Similarity: 80.682 Percent Identity: 57.955
alignment_block:
us-09-898-586-24 x BH111304 ..
Align seg 1/1 to: BH111304 from: 1 to: 788
34 TyrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLe 50
|||||
6 TACATCTTCAGCCCTGCTCGCAATGCAATTAATCTTAGTACATCTGCCCT 55
|||||
50 uasPserArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaV 67
|||||
56 GGACCCCAAGCTGCACACCCCTATGTACTTCTTCTTCATCTGGCCA 105
|||||
67 alValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 83
|||||
106 TTTCTGTACATATCTATGCTGCCAACAATGTCCCAAGATGTTGCCAAC 155
|||||
84 LeuLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnTh 100
|||||
156 CTATATAACCGAAGAAAGAACCATCTCATTTGGCCCATGCATCACACAGAC 205
|||||
100 rPheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValM 117
|||||
206 ATCTTGTGATTTGGCCCTTGGCGCTTGGAGTCTTGATTTGGCAGCCA 255
|||||
117 etSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAl 133
|||||
256 TGTCTTATACACAGGTTTGTGGCCATCTGCCACCCCTACACTACACTGT 305
|||||
133 alIleMetThrTrpArgValAlaCysIleThrLeuAlaValThrSerTrpTh 150
|||||
306 CATCATGAGTTGGAAAGATATGTGTGGCCCTGGCTGTACATCTGTGCAT 355
|||||
150 hrGlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuPro 166
|||||
356 GTGATTTAGCCCTCTCTGTGGACACACACAATTCCTGCTAAGGTTGCC 405
|||||
167 PheCysArgProGlnIleTyrHisPhePheCysGluIleLeuAlaVal 183
|||||
406 TTTGTGGGCCCCAGGAAATTAACCACTCTTCTGTGAAATCCTGCTGT 455
|||||
183 lLeuLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuA 200
|||||
456 TCTCAAGCTGGCCCTGCGCCGACACTTGTATCAACCAATCGTCATCTAG 505
|||||
200 laGlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSe 216
|||||
506 CTGCATGTGTTTCTGCTTAGTTGGACCCCTTGTCTAGTGTGTATC 555
|||||
216 rTyrMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValG 233
|||||
556 TTACACGACATCTCTGGACAATCTGAAATGCAAGTCAAAAGAGGCTC 605
|||||
233 lArgIlyAlaPheCysThrCysPheSerHisIleCysValIleGlyLeu 249
|||||
606 GCAGGAAGGCTTTTCCACCTGTCTCCATCTCTGCTGTGTGTGGGCTC 655
|||||
250 PheTyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsn 265
|||||
656 TTCTTTGGTATAGCCATGTGTGTTATATATGTCCTGACTCTGATCAG 705
|||||
266 ProLysGluClnLysTyrLeuLeuLeuPheHisSerLeuPheAsnPr 282
|||||
706 CGAGAGGACGAGGAAATTTGTCACTGTTCACAGTCTTTTACCCA 755
```



```
282 OMeLeuAsnProLeuIleCysSerLeuArgAsn 293
      |||
      |||
756 AAGTGGATTCCTCATTAAT...AGTCTGAGGAAC 786
seq_name: gb_est2:BF974217

seq_documentation_block:
LOCUS      BF974217                671 bp      mRNA      linear      EST 22-JAN-2001
DEFINITION  602243851F1 NIH_MGC_48 Homo sapiens cDNA clone IMAGE:4334965 5',
            mRNA sequence.
ACCESSION   BF974217
VERSION     BF974217.1  GI:12341432
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 671)
AUTHORS    NIH-MGC http://mgc.ncl.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
            cDNA Library Preparation: Ling Hong/Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            plate: LICM1205 row: e column: 14
            High quality sequence stop: 666.
            Location/Qualifiers
                1..671
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:4334965"
                /clone_lib="NIH_MGC_48"
                /tissue_type="primary B-cells from tonsils (cell line)"
                /lab_host="DH10B (phage-resistant)"
                /note="Organ: B-cells; Vector: pOTB7; Site_1: XhoI;
                Site_2: EcoRI; cDNA made by oligo-dT priming.
                Directionally cloned into EcoRI/XhoI sites using the
                following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                for average insert size 1.8kb. Library constructed by Ling
                Hong in the laboratory of Gerald M. Rubin (University of
                California, Berkeley) using ZAP-cDNA synthesis kit
                (Stratagene) and Superscript II RT (Life Technologies).
                Note: this is a NIH-MGC Library."
BASE COUNT      184 a      146 c      219 g      122 t
ORIGIN
alignment_scores:
      Quality: 687.00      Length: 226
      Ratio: 3.734      Gaps: 7
Percent Similarity: 81.416      Percent Identity: 65.487

alignment_block:
US-09-898-586-24 x BF974217/rev ..
Align seg 1/1 to reverse of: BF974217 from: 1 to: 671
58 MetTyrPhePheLeuSerHisLeuAlaValAlaSerPheAlaTyrAlaCys 74
|||||
671 ATGTACTTCTTCTCTACACCTGGCGCTGTCACATCGCCTATGCTG 622
74 sAsnThrValProArgMetLeuValAsnLeuLeuHisProAlaLysProI 91
|||||
621 CAACACAGTGGCCAGATGCTGTGAACCTCTGCATCCACCAAGCCCA 572
91 leSerPheAlaGly...ArgMetMetGlnThrPheLeuPheSerThrPhe 106
|||||
```

```
571 TCTCCTTTGGCTGGCATGATGACAGTAGACCTTTCTCTTGTGCTGACTTTT 522
107 AlaValThrGluCysLeuLeuLeuValValMetSerTyrAspLeuTyrVa 123
||| |||
521 GCACATACTGAATGACCTCCTCTGTGGTGTGCTGATGTCTTAGATCGGTACGT 472
123 lAlaIleCysHisProLeuArgTyrLeuAlaIleMetThrTrpArgValC 140
|||||
471 GGCCATCTGGCCACCCTCTCCGATATTTCATCATCATGACCTGGAAAGTCT 422
140 ysIleThrLeuAlaValThrSerTrpThrThrGlyValLeuLeuSerLeu 156
|||||
421 GCATCACTCTGGCCATCATCTCTGGACATGTGGCTCTCCTGCTATG 372
157 lIleHisLeuValLeuLeuLeuProLeuProPheCysArgProGlnLysI 173
:::
371 GTCCATGTGAGCCCTCATCTCTAAGACTGCCCTTTGTGGGCTCGTGAAT 322
173 eTyrHisPhePheCysGluIleLeuAlaValLeuLysLeuAlaCysAl 189
| |||
321 CAACCACTTCTTCTGTGAATCTCTGTCTGCTCAGAGCCCTAGGCCCTGTGC 272
189 aAspThrHisIleAsnGlu...AsnMetValLeuAlaGlyAlaIleSerG 205
|||||
271 TGATACCATGAGCTCAACCAAGGTAGTCACTTTTGACAGCCTGCATGTCAAT 222
205 lYleuValGlyProLeuSerThrIleValValSerTyrMetCysIleLe 221
|||||
221 ACCTGTAGGAGCACTCTGCTGCTGTGCTGTCTCTCTACTCACAACATCTCT 172
221 u.CysAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysAlaPhe 237
|:::
171 AGCGGGCCATCTCTGAGATCCAGTCTGGGGAGAGGCCGCAAGCCCTTTC 122
238 CysThr.CysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThra 254
||| |||
121 TCCACCATGCTCTCTCCACCTCTGCGTAGTGGACTCTTCTTTGGCAGCG 72
254 lAlaIleMetTyrValGlyProArgTyrGlyAsnProLysGluGlnLys 270
|||||
71 CCATGCTCATGTACATGGCCCTTAAGTCCCGCATCTCTGAGGAGCAGCAG 22
271 LysTyrLeuLeuLeuPhe 276
||| |||
21 AAGGTCCTTTTCTATT 4

seq_name: gb_est2:BF116115

seq_documentation_block:
LOCUS      BF116115                732 bp      mRNA      linear      EST 24-OCT-2000
DEFINITION  7n77g04.x1 NCI-CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3570631 3',
            similar to TR:095047 095047 WUGSC:H_DJ0988G15.2 PROTEIN.;, mRNA
            sequence.
ACCESSION   BF116115
VERSION     BF116115.1  GI:10985591
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 732)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento
            Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
```

found through the I.M.A.G.E. Consortium/LLNL, send email to:  
info@image.llnl.gov  
Seq primer: -40UP from Gibco  
High quality sequence stop: 472.

## FEATURES

## source

1..732  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="NCI\_CGAP\_Ov18"  
/tissue\_type="fibrotheoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: ovary; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAGTGGGAGCGCGCGGCACATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

## BASE COUNT

152 a 217 c 145 g 218 t

## ORIGIN

## alignment\_scores:

Quality: 678.00 Length: 172  
Ratio: 4.431 Gaps: 1  
Percent Similarity: 88.953 Percent Identity: 79.070

## alignment\_block:

US-09-898-586-24 x BF116115 ..

Align seg 1/1 to: BF116115 from: 1 to: 732

```

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||:|||||  ||| :|||:|||||:|||||:|||||:|||||
219 ATGGGGGAAATCAGACAAATGTCACAGAGTCTCTCTACTGGGATTTC 268
17 oval.GlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 33
::: |||||:|||||:|||||:|||||:|||||:|||||:|||||
269 CTTGGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCCCTGTTTC 318
34 TyrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLe 50
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
319 TATATCTTCACTGCTGGGGGAACGGGGCCATCTGGGGTCATCTCTCACT 368
50 uAspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaV 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||
369 GGACTCCAGACTCCACACCCCATGTACTTCTTCTCTCACACCTGGCTG 418
67 aValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||
419 TCGTGCACATCGCTACACCCGCCAACACAGGTGCCCCAGATCTGGCGAAC 468
84 LeuLeuHisProAlaIysProIleSerPheAlaGlyArgMetMetGlnTh 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||
469 CTCTGCATCCAGCAAGCCCATCTCTTGGCTGCTGATGACGACGAC 518
100 rPheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValM 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||
519 CTTTCTCTGTTGAGTTTGGACACAGCAATGTCTCTCTGCTGCTGA 568
117 eTserTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||
569 TGTCTTACGATCGTTACGTGGCCATCTGCCACCCCTCGATACCTCCGTC 618
134 IleMetThrTyrArgValCysIleThrLeuAlaValThrSerTyrThrTh 150
|||||:|||||:|||||:|||||:|||||:|||||:|||||
619 ATCATGACCTGGAGAGTCTGCATCACCTGGCCGTCACCTTCCCTGACGTG 668
150 rGlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProp 167
||| |||||:|||||:|||||:|||||:|||||:|||||

```

669 TGGCTCCCTCTGCTCTGGCCCATGTGTTCTCATCTAAGACTGCCCT 718

167 heCysArgProGln 171  
|||

719 TCTCTGGGCTTCAA 732

seq\_name: gb\_est1:AI340119

seq\_documentation\_block:

LOCUS AI340119 367 bp mRNA linear EST 29-DEC-1998

DEFINITION qk64f08.x1 NCI\_CGAP\_Co8 Homo sapiens cDNA clone IMAGE:1873767 3' similar to SW:OLF7\_MOUSE P34984 OLFATORY RECEPTOR-LIKE PROTEIN R7 ;, mRNA sequence.

ACCESSION AI340119 GI:4077046

VERSION AI340119.1 GI:4077046

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 367)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgaps-remail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco

High quality sequence stop: 245.

location/Qualifiers

1..367

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:1873767"

/clone\_lib="NCI\_CGAP\_Co8"

/tissue\_type="adenocarcinoma"

/lab\_host="DH10B"

/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 115 a 73 c 116 g 63 t

ORIGIN

## alignment\_scores:

Quality: 643.00 Length: 122  
Ratio: 5.270 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-898-586-24 x AI340119/rev ..

Align seg 1/1 to reverse of: AI340119 from: 1 to: 367

```

60 PhePheLeuSerHisLeuAlaValAlaAspIleAlaTyrAlaCysAsnTh 76
|||||:|||||:|||||:|||||:|||||:|||||:|||||
367 TTCTTCTCTACACCTGGCGGTCTGACATCGCTACGCTGACACAC 318
76 rValProArgMetLeuValAsnLeuLeuHisProAlaIysProIleSerP 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

317 GGTGCCCCGGATGCTGTGAACCTCTCGATCCAGCCAGCCCATCTCT 268
93 heAlaGlyArgMetMetGlnThrPheLeuPheSerThrPheAlaValThr 109
|||||
267 TTGCGGGCCGCATGATGACGACCTTCTGTGTTTCCACTTTTGCTGTCA 218
110 GluCysLeuLeuValValMetSerTyrAspLeuTyrValAlaIleCys 126
|||||
217 GAATGTCTCTCTCGTGTGTGATGTCTCTATGATCTGTACGTGGCCATCTG 168
126 SHisProLeuArgTyrLeuAlaIleMetThrTyrPArgValCysIleThrL 143
|||||
167 CCACCCCTCCGATATTGTGGCCATCATGACCTGGAGAGTCTGCATCACCC 118
143 euAlaValThrSerTyrThrThrGlyValLeuLeuSerLeuIleHisLeu 159
|||||
117 TCGCGGTGACTTCTCTGGACCACTGAGTCTTTTATCTGTGATTCATCTT 68
160 ValLeuLeuLeuProLeuProPheCysArgProGlnIleTyrHisPhe 176
|||||
67 GTGTACTTCTACCTTACCTTCTGTAGGCCCCAGAAATTTATCACTT 18
176 ePheCysGluIleLeu 181
|||||
17 TTTTGTGAATCTTA 2

seq_name: gb_est1:AI285267

seq_documentation_block:
LOCUS AI285267 361 bp mRNA linear EST 29-JAN-1999
DEFINITION q138c08.x1 NCI-CGAP_C08 Homo sapiens cDNA clone IMAGE:1874606 3'
similar to SW:OLF7_MOUSE P34984 OLFACTOR RECEPTOR-LIKE PROTEIN K7
; mRNA sequence.
ACCESSION AI285267
VERSION AI285267.1 GI:3923500
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 361)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 659 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 26.
FEATURES
source
location/Qualifiers
1..361
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1874606"
/clone_1lb="NCI-CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
colon adenocarcinoma, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by

```

```

BASE COUNT 115 a 73 c 110 g 63 t
ORIGIN
alignment_scores:
Quality: 602.00 Length: 120
Ratio: 5.190 Gaps: 0
Percent Similarity: 96.667 Percent Identity: 95.833
alignment_block:
US-09-898-586-24 x AI285267/rev ..
Align seg 1/1 to reverse of: AI285267 from: 1 to: 361
60 PhePheLeuSerHisLeuAlaValAlaAspIleAlaTyrAlaCysAsnTh 76
|||||
361 TTTTCCCTTCACACCTGGCGGTCTTGACATTTGCCCTAGCCCTGCACAC 312
76 rValProArgMetLeuValAsnLeuLeuHisProAlaIleCysProIleSerP 93
|||||
311 GGTGCCCCGGATGCTGTGAACCTCTGCATCCAGCCAGCCCATCTCTCT 262
93 heAlaGlyArgMetMetGlnThrPheLeuPheSerThrPheAlaValThr 109
|||||
261 TTGCGGGCCGCATGATGACGACCTTCTGTTCACCTTTGCTGTCA 212
110 GluCysLeuLeuValValMetSerTyrAspLeuTyrValAlaIleCys 126
|||||
211 GAATGTTCCTCTCGTGTGAAGTCCAAATGATTTGTAGTGGCCATCTG 162
126 SHisProLeuArgTyrLeuAlaIleMetThrTyrPArgValCysIleThrL 143
|||||
161 CCACCCCTCCGATATTGTGGCCATCATGACCTGGAGAGTCTGCATCACCC 112
143 euAlaValThrSerTyrThrThrGlyValLeuLeuSerLeuIleHisLeu 159
|||||
111 TCGCGGTGACTTCTCTGGACCACTGAGTCTTTTATCTGTGATTCATCTT 62
160 ValLeuLeuLeuProLeuProPheCysArgProGlnIleTyrHisPhe 176
|||||
61 GTGTACTTCTACCTTACCTTCTGTAGGCCCCAGAAATTTATCACTT 12
176 ePheCysGlu 179
|||||
11 TTTTGTGA 2

seq_name: gb_est2:H62441

seq_documentation_block:
LOCUS H62441 426 bp mRNA linear EST 06-OCT-1995
DEFINITION yr79d08.r1 Soares fetal liver spleen INFUS Homo sapiens cDNA clone
IMAGE:211503 5' similar to SP:OLF7_MOUSE P34984 OLFACTOR
RECEPTOR-LIKE PROTEIN K7 ; mRNA sequence.
ACCESSION H62441
VERSION H62441.1 GI:1015273
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 426)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W., Hawkins
M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N., Mardis, E., Moore
B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T.,
Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E.,
Underwood, R., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine

```

Fri Aug 23 16:18:05 2002

us-09-898-586-24.rst

Page 12

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estewatson.wustl.edu  
Insert Size: 1867

119 yrAspleuTyRvaIAlaIecyShisproLeuArgTyR 131  
 |||||  
 350 ATGATCTGTACGTNGCCATCTGNCACCCCTCCGATAT 388

High quality sequence stops: 304  
Source: IMAGE Consortium, LINL.  
This clone is available royalty-free through LINL ; contact the  
IMAGE Consortium ([info@image.linl.gov](mailto:info@image.linl.gov)) for further information.  
Insert Length: 1867      Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 304.

## FEATURES

**URES**  
**source**

```

1. 426
/organism="Homo sapiens"
/db_xref="GDB:3784344"
/db_xref="taxon:9606"
/clone="IMAGE:211503"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/notice="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - Oligo(dT) primer
[5' AACTGGAGAGATTAAATTAAGATCTTTTCTTTTCTTTTCTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaudo."
a 140 c 95 g 111 t 5 others

```

BASE COUNT  
ORIGIN

75 a	140 c	95 g	111 t	5 others
------	-------	------	-------	----------

alignment\_scores:

Quality:	593.00	Length:	131
Ratio:	4.669	Gaps:	2
Similarity:	96.947	Percent Identity:	96.947

alignment\_block:

586-24 X H62441 . .

Align seg 1/1 to: H62441 from: 1 to: 426

3	AspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheProValG	19
1	GACATATATACATCCATCACAGAGTTCCTCCTACTGGGATTTCCCGTTGG	50
19	yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyValP	36
51	CCCAAGGATTCAAGTGGCTCCTTTGGGCTCTTCCCTGTCTTACGTC	100
36	heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer	52
101	TCACCCCTGCTGGGGAAACGGGACCATACTGGGGCTCATCTCGACTCC	150
53	ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValVala	69
151	AGACTGCACGCGCCCATGTACTCTTCCTCTCACACACTGGGGCGGTG	200
69	spIleAlaTyraIaCysAsnThrValProArgMetLeuValAsnLeuLeu	85
201	ACATCGCCTACGCGCTGCACACACGGTGCCTCCGGATGCTGGTGAACCTCTCTG	250
86	HisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLe	102
251	CATCCAGGCCAAGCCCATCTCTCTTTGGGGGCGGATGATGCAGACCTTTCT	299
102	upheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSer	119
300	GTTTTCACACTTTGGCTGTACACAGATGTCTCTCCGGGTGGTGATGTCT	349

OM of: US-09-898-586-24 to: Issued\_Patents\_NA:\* out\_format: pfs

Date: Aug 21, 2002 8:19 PM

About: Results were produced by the GenCore software, version 4.5,  
Copyright (c) 1993-2000 CompuGen Ltd.

#### Command line parameters:

```
-MODEL=framet+pn.model -DEV=xlh
-Q=/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-467-948A-1 + 1048.50 2174.83 1.9e-113 1713 1
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09898586_@CGN1_1_48 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPHY -WAIT -THREADS=1
```

#### Search information block:

Query: US-09-898-586-24  
Query length: 310  
Database: Issued\_Patents\_NA:\*  
Database sequences: 383533  
Database length: 122816752  
Search time (sec): 45.940000

Sequence	Strd Orig	ZScore	EScore	Len	Documentation	..	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961	962	963	964	965	966	967	968	969	970	971	972	973	974	975	976	977	978	979	980	981	982	983	984	985	986	987	988	989	990	991	992	993	994	995	996	997	998	999	1000	1001	1002	1003	1004	1005	1006	1007	1008	1009	1010	1011	1012	1013	1014	1015	1016	1017	1018	1019	1020	1021	1022	1023	1024	1025	1026	1027	1028	1029	1030	1031	1032	1033	1034	1035	1036	1037	1038	1039	1040	1041	1042	1043	1044	1045	1046	1047	1048	1049	1050	1051	1052	1053	1054	1055	1056	1057	1058	1059	1060	1061	1062	1063	1064	1065	1066	1067	1068	1069	1070	1071	1072	1073	1074	1075	1076	1077	1078	1079	1080	1081	1082	1083	1084	1085	1086	1087	1088	1089	1090	1091	1092	1093	1094	1095	1096	1097	1098	1099	1100	1101	1102	1103	1104	1105	1106	1107	1108	1109	1110	1111	1112	1113	1114	1115	1116	1117	1118	1119	1120	1121	1122	1123	1124	1125	1126	1127	1128	1129	1130	1131	1132	1133	1134	1135	1136	1137	1138	1139	1140	1141	1142	1143	1144	1145	1146	1147	1148	1149	1150	1151	1152	1153	1154	1155	1156	1157	1158	1159	1160	1161	1162	1163	1164	1165	1166	1167	1168	1169	1170	1171	1172	1173	1174	1175	1176	1177	1178	1179	1180	1181	1182	1183	1184	1185	1186	1187	1188	1189	1190	1191	1192	1193	1194	1195	1196	1197	1198	1199	1200	1201	1202	1203	1204	1205	1206	1207	1208	1209	1210	1211	1212	1213	1214	1215	1216	1217	1218	1219	1220	1221	1222	1223	1224	1225	1226	1227	1228	1229	1230	1231	1232	1233	1234	1235	1236	1237	1238	1239	1240	1241	1242	1243	1244	1245	1246	1247	1248	1249	1250	1251	1252	1253	1254	1255	1256	1257	1258	1259	1260	1261	1262	1263	1264	1265	1266	1267	1268	1269	1270	1271	1272	1273	1274	1275	1276	1277	1278	1279	1280	1281	1282	1283	1284	1285	1286	1287	1288	1289	1290	1291	1292	1293	1294	1295	1296	1297	1298	1299	1300	1301	1302	1303	1304	1305	1306	1307	1308	1309	1310	1311	1312	1313	1314	1315	1316	1317	1318	1319	1320	1321	1322	1323	1324	1325	1326	1327	1328	1329	1330	1331	1332	1333	1334	1335	1336	1337	1338	1339	1340	1341	1342	1343	1344	1345	1346	1347	1348	1349	1350	1351	1352	1353	1354	1355	1356	1357	1358	1359	1360	1361	1362	1363	1364	1365	1366	1367	1368	1369	1370	1371	1372	1373	1374	1375	1376	1377	1378	1379	1380	1381	1382	1383	1384</
----------	-----------	--------	--------	-----	---------------	----	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	--------



```
128 ATGGTGAATAATCAGACAATGGTCACAGAGTTCCCTCTACTGGGATTTCT 177
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhePhe 34
:::|||||
178 CTTGGGCCAAGGATTCAGATGCTCTCTTGGGCTCTTCTCTCCCTGTTCT 227
34 yrvAlPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
|||||
228 ATGCTTCACCCCTGCTGGGGAATGGGACCAATCTGGGGCTCATCTCAGCTG 277
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
|||||
278 GACTCCAGACTCCACACCCCATGTACTCTTCTCTCTCACACCTGGCCGT 327
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
|||:::|||||
328 CGTCAACATCGCCTATGCTGCACACAGAGTGCCTCCAGATGCTGGTAACC 377
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
378 TCCTGCATCCAGCCCAAGCCCAATCTCTTGTGCTGTCATGACACTAGAC 427
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValAlme 117
|||||
428 TTTCTCTTTTGTAGTTTGGACATACTGAATGCCCTCCTGTGTGTCGAT 477
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrIleAlaI 134
|||||
478 GTCTACGATCGGTACGTGGCCATCTGCACACCCCTCCGATATTTCATCA 527
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150
|||||
528 TCATGACCTGGAAAGTCTGCATCATCTGGGCAATCCTTCCCTGGACATGT 577
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
||| |||:::|||||
578 GGCTCCCTCCTGGCTATGCTCATGTGAGCCCTCATCTTAAGACTGCCCTT 627
167 ecysArgProGlnIleTyrHisPhePheCysGluIleLeuAlaValL 184
||| |||:::|||||
628 TTGTGGGCTCGTGAATCAACCACTTCTCTGTGAATCCTGTCTGTCC 677
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||:::|||||
678 TCAGGCTGGCTGTGCTGATACCTGCTCAACCAAGTGTGTCATCTTGAA 727
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
:::|||||
728 GCCTGCATGTTTCATCCTGTGGGACCACTCTGCTGGTGTCTCTCTA 777
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
| ||| |||:::|||||
778 CTCACACATCCTGGGGGCATCCTGAGGATCCAGTCTGGGGAGGGCCGA 827
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
|||||
828 GAAAGGCTCTCTCCACCTGCTCTCCACCTCTGCTGAGTGGACTCTTC 877
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
:::|||||
878 TTTGGSAGCGCATGCTCATGTACATGGCCCTTAAGTCCCGCCATCTGA 927
267 sGluGlnLysIleTyrLeuLeuLeuPheHisSerLeuPheAsnPro... 282
:::|||||
928 GGAGCAGCAGAAAGTCTTTTCTTATTATTACAGTTCCTTTCACACCCCGA 977
283 MetLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsn 298
|||||
978 TGCTTAACCCCTGATTTACACACCTGAGGAATGTAGAGGGTCAAGGGT 1027
299 ThrLeuLys 301
:::|||||
1028 GCCCTCCGA 1036
```

```
seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-467-947A-1
seq_documentation_block:
; Sequence 1, Application US/08467947A
; Patent No. 6090575
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; TITLE OF INVENTION: Coupled Receptor GPR1
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,947A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2600
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1713 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 116..1003
US-08-467-947A-1

alignment_scores:
Quality: 1048.50 Length: 305
Ratio: 4.017 Gaps: 4
Percent Similarity: 85.574 Percent Identity: 69.180

alignment_block:
US-09-898-586-24 x US-08-467-947A-1 ..
Align seg 1/1 to: US-08-467-947A-1 from: 1 to: 1713

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
||| ||| |||:::|||||
128 ATGGTGAATAATCAGACAATGGTCACAGAGTTCCCTCTACTGGGATTTCT 177
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhePhe 34
:::|||||
178 CTTGGGCCAAGGATTCAGATGCTCTCTTGGGCTCTTCTCTCCCTGTTCT 227
```

34 yrvAlPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
228 ATGTCTTCAACCTGCTGGGGAATGGGACCATCTGGGGCTCATCTCACTG 277  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67  
|||||  
278 GACTCCAGACTCCACACCCCATGTACTTCTCTCTCTCACACCTGGCCGT 327  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
|||||  
328 CGTCAACATCGCCCTATGCTTGCACACAGATGCCCCAGATGCTGTGAACC 377  
84 euLeuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
378 TCCTGCATCCAGCCAGCCCATCTCTGCTGCTGTCATGACACTAGAC 427  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
428 TTTCTCTTTTGAAGTTTGGACATAGTATGCTCTCTGCTGTGCTGCTGAT 477  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
478 GTCCCTACGATCGGTACGTGGCCATCTGCCACCCCTCCGATATTTCATCA 527  
134 lMetThrTriParGValCysIleThrIleAlaValThrSerTrpThrThr 150  
|||||  
528 TCATGACCTGGAAGTCTGCATCAGCTGGGCATCACTCTCCGAGACATGT 577  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167  
|||||  
578 GGCTCCCTCCTGGCTATGGTCCATGTGAGCCTCATCTTAAGACTGCCCTT 627  
167 ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184  
|||||  
628 TTGTGGGCTCGTGAATCAACACACTTCTTGTGAATCCTGTCTGTCC 677  
184 euLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
678 TCAGGCTGGCTGTGCTGATACCTGGCTCAACAGGTGTCATCTTTGAA 727  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
|||||  
728 GCCTGCATGTTCATCTGCTGGGACCACTCTGCCTGCTGCTGTCTCTTA 777  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgIuValGlnA 234  
|||||  
778 CTCACACATCCTGGGGGCATCTGAGATCCAGTCTGGGGAGGGCCGCA 827  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
828 GAAAGGCTTCTCCACCTGCTCTCCACCTCTGCGTAGTGGGACTCTTC 877  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
|||||  
878 TTTGGSAGCGCATCGTATGTATGATGGCCCTAAGTCCCGCATCTCTGA 927  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPro... 282  
|||||  
928 GGAGCAGCAGAAAGTCTTTTCTTATTTACAGTTCCTTCAACCCCGA 977  
283 MetLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsn 298  
|||||  
978 TGCTTAAACCCCTGATTTTACAAACCCGAGGAATGTAGAGGGTCAAGGGT 1027  
299 ThrLeuLys 301  
|||||  
1028 GCCCTCCGA 1036

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-748-506-5  
seq\_documentation\_block:  
; Sequence 5, Application US/08748506

Patent No. 6159707  
GENERAL INFORMATION:  
APPLICANT: Ronnett et al.  
TITLE OF INVENTION: NOVEL SPERM RECEPTORS  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leydig, Volt & Mayer, Ltd.  
STREET: Two Prudential Plaza, Suite 4900  
CITY: Chicago  
STATE: IL  
COUNTRY: US  
ZIP: 60601-6780  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/748,506  
FILING DATE: 08-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/033,751  
FILING DATE: 09-NOV-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
REFERENCE/DOCKET NUMBER: 74940  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-616-5600  
TELEFAX: 312-616-5700  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 966 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-748-506-5

alignment\_scores:  
Quality: 609.00 Length: 308  
Ratio: 2.819 Gaps: 0  
Percent Similarity: 70.130 Percent Identity: 40.909

alignment\_block:  
US-09-898-586-24 x US-08-748-506-5 ..  
Align seg 1/1 to: US-08-748-506-5 from: 1 to: 966

3 AspAsnIleThrSerIleThrGluPheLeuLeuGlyPheProValG1 19  
:::|||||  
28 GAGAAATAGTTGACTGTCAACACCTTGGCATTTGCCAAGTTCTCTGAGGT 77  
19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValP 36  
|||  
78 CCCTGGAGAAATGCTTCCCTCTGTTCAACCTCAACCTTCATGTTCTTAG 127  
36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52  
:::|||||  
128 TATCACTAACAGGAATACTCTCATAGTCTTGTCTTATTTGACCACTGCA 177  
53 ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValAlas 69  
|||||  
178 TCTCTACACACCCCATGTACTTCTTCTGCGCACTGTCTCTCTGGA 227  
69 pIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86  
:::|||||  
228 GATTTGGCTTAACTTGTCTGTGTCATACCCCAAGATGCTGCAGAGCCTTGTGA 277  
86 isProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102  
|||||  
278 GTGAGGCCAGAGAGATCTCTCGGAGGAGGATGTGCCACACAGATGTTTTP 327

```
103 pheSerThrPheAlaValThrGluCysLeuLeuValValMetSerTy 119
||||: |||:||||| |||:|||||:||||:
328 TTCGATTTTGTGTTAACTAGTGTGCTGCTATTGGCAGCCATGGCTT 377
119 rAsPleuTyValAlaIleCysHisProLeuArgTyLeuAlaIleMet 136
||| :|||:||||| |||:||||| :||| :||| :
378 TGACCGCTGCATGGCCATATGTCCCACTCCACTATGCACCCGATGA 427
136 hTrtpArgValCysIleThrLeuAlaValThrSerTrpThrGlyVal 152
: :|||:||||| |||:||||| :||| :||| :
428 CTCGNGAGGTATGTGCCCATTTGGCAATGTGTTCATGGGGAATGGATGC 477
153 LeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPheCysAr 169
: :|||:||||| :||| :||| :||| :
478 ATAGTAAGTCTGGGACAAACCAATTTTATTTCTCCTTGAACCTCTGTGG 527
169 gProGlnLysIleTyHisPhePheCysGluIleLeuAlaValLeuLysL 186
||| :|||:||||| |||:||||| :||| :||| :
528 ACCCTGTGAATATAGACACACTTCTGTGTGACCTTCACCTCTCTGGCAC 577
186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAla 202
|||||:||||| |||:||||| :||| :||| :
578 TTGCCTGTGGAGATACATCCCAAAACGAGGCTGCCAATCTTGTGTAGCA 627
203 IleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrmety 219
: :|||:||||| |||:||||| :||| :||| :
628 GTCTCTGCATATCTAGCCCATTTTGTGCTGATCATTTATTTATGTCAA 677
219 sIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysA 236
||||| |||:||||| :||| :||| :||| :
678 AATTCTCATGTGAGTGTCTGCTGATGCTTCACTGAGGCGGCATTAAG 727
236 lAphCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyrgly 252
|||:||||| |||:||||| |||:||||| :||| :||| :
728 CTCTTTCACCTGTGTCTGTCTCCTACTGTGTAGTCACACTTTTATGGC 777
253 ThrAlaIleIleMetTyValGlyProArgTyrglyAsnProLysGluG 269
: :|||:||||| |||:||||| :||| :||| :
778 TCAGCATGTATTAACCTATTGTGAGGCCCAAGTCTAGCCACTCACAGGAAT 827
269 nLysLysTyrlLeuLeuPheHisSerLeuPheAsnProMetLeuAsn 286
: :|||:||||| |||:||||| :||| :||| :
828 GGACAAATCTTGGCCCTCTTCTACACAGTAGTACATCCATGCTGAACC 877
286 roLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArg 302
|||:||||| |||:||||| |||:||||| :||| :||| :
878 CTATCATCTATAGTTTAAAGAACAGGAAGTCAAGCGACACTGAGAGA 927
303 ValLeuGlyValGluArgAlaLeu 310
: :|||:||||| |||:||||| :||| :||| :
928 ACTCTGGCGCTGAAAAAATTTCTG 951
```

seq\_name: /cgn2\_6/plodata/2/ina/5B\_COMB.seq:US-08-827-291A-1

```
seq_documentation_block:
; Sequence 1, Application US/08827291A
; Patent No. 5874243
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Sathe, Ganesh
; TITLE OF INVENTION: NOVEL OLRCC15 RECEPTOR
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/827,291A
;; FILING DATE: 28-MAR-1997
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: King, William T
;; REGISTRATION NUMBER: 30,954
;; REFERENCE/DOCKET NUMBER: GP50001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 610-270-5015
;; TELEFAX: 610-270-5090
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1290 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; US-08-827-291A-1
```

```
alignment_scores:
      quality: 606.00      length: 304
      Ratio: 2.742      Gaps: 2
Percent Similarity: 72.697      Percent Identity: 40.132
```

alignment\_block:

US-09-898-586-24 x US-08-827-291A-1 ..

Align seg 1/1 to: US-08-827-291A-1 from: 1 to: 1290

```
3 AspAsnIleThrSerIleThrGluPheLeuLeuGly..PheProva 18
: :|||:||||| :|||:|||||:||||:
305 GAGATTCGACCTTCACACTCCGACTTCATCTTCCGGAATCTCAATCA 354
18 lGlyProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheTyrv 35
: :|||:||||| :|||:||||| :|||:||||| :
355 CAGCCCAACCCACACACCTTCTCTTCTGTGCTCTGCGCATCTTTTCAG 404
35 alphaThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAsp 51
|| :|||:||||| :|||:||||| :|||:||||| :
405 TG..GCCTTCATGGGAACCTCTGTATGCTTCTCCTCATCTACCTGGAC 451
52 SerArgLeuHisAlaProMetTyrrPhePheLeuSerHisLeuAlaVal 68
: :|||:||||| :|||:||||| :|||:||||| :
452 ACCCAGCTCCACACCCCACTGTACCTCTCTCAGCCCACTGTCCCTCAT 501
68 lAspIleAlaTyAlaCysAsnThrValProArgMetLeuValAsnLeuL 85
: :|||:||||| :|||:||||| :|||:||||| :
502 GGACCTCATGTCTCATCTGCACCAACGATGACCAAGATGGCTTCACAC 551
85 euHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPhe 101
|| :|||:||||| :|||:||||| :|||:||||| :
552 TGCTGGCAGCAAGTCCATTTCTATGGCTGTGTGCGCACACAAATTTTC 601
102 LeuPheSerThrPheAlaValThrGluCysLeuLeuValValMetSe 118
: :|||:||||| :|||:||||| :|||:||||| :
602 TTCTATACATCACTGCTGGCTCTGAAATGCTTTCTTTTGGCTGTATGGC 651
118 rTyAsPleuTyValAlaIleCysHisProLeuArgTyLeuAlaIleM 135
: :|||:||||| :|||:||||| :|||:||||| :
652 TTATGACCGCTACACTGCCATTTGCCACCCCTCAAGATACACCAATCTCA 701
135 eThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrGly 151
|||:||||| :|||:||||| :|||:||||| :
702 TGAGCCCTAAATTTGTGGAATTATGACTGCTTTTCTCTGAAATCTTGGGC 751
```

```

152 ValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPheCy 168
      ::::::::::: ||| :::: :::: ::::|
752 TCTACAGATGGAATCATTTATGCTGTAGCCACATTTTCTCTCTCTACTG 801
168 sargProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuL 185
      ::::::::::: ||| ||||| ||||| ::::|
802 TGGTCTCGGAAATAGCCCACTTCTTCTGTGAGTTACCTTCCCTACTAA 851
185 ysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAgl 201
      |||::||| ||||| ||| |||:::|
852 TCCTCTCATGCAATGACACATCATATTGAAAAGGTATTTCATTTC 901
202 AlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMe 218
      ::||| ||||| |||:::| |||:::| |||:::|
902 TCTATAGTAAATGCTTGTGTTTCCCTGTGCAATCATCATTTGCTTCTATGC 951
218 tCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnArgL 235
      ::::: |||:::| ||| ||| |||:::|
952 TGGAGTTATTCTGGCTGTTCATTCACATGGGATCTGGAGAGGGTCTCGCA 1001
235 ysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyr 251
      ||||| ||||| ||||| ||||| |||:::|
1002 AAGCTTTCACGACCTGTCTCTCTCACCCTCATGTGTGGGAAATGTTCTAT 1051
252 GlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLysG 268
      |||:::| |||:::| |||:::| |||:::|
1052 GGAGCAGGTTGTTCATGTACATACAGCCACATCTGATCGCTCCCAAC 1101
268 uGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuA 285
      ||| ||| ::::| |||:::| |||:::| |||:::|
1102 GCAGACAAAGCTGTGTCTGTATTTCTACACCATCTCCTCCTCCATGCTGA 1151
285 snProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLys 301
      ||||| ||||| ||||| ||||| |||:::|
1152 ATCCCTCATCTACAGCCCTCCGCAACAAGAGACGACGATTCATG 1201
302 ArgValLeuGly 305
      ::::: |||
1202 AAGATCTCAGGA 1213

```

seq\_name: /cgn2\_6/ptodata/2/lna/6A\_COMB.seq:US-08-748-506-7

seq\_documentation\_block:

; Sequence 7, Application US/08748506  
; Patent No. 6159707

; GENERAL INFORMATION:

; APPLICANT: Ronnet et al.

; TITLE OF INVENTION: NOVEL SPERM RECEPTORS

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leydig, Volt & Mayer, Ltd.

; STREET: Two Prudential Plaza, Suite 4900

; CITY: Chicago

; STATE: IL

; COUNTRY: US

; ZIP: 60601-6780

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/748,506

; FILING DATE: 08-NOV-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/033,751

; FILING DATE: 09-NOV-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; REFERENCE/DOCKET NUMBER: 74940

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-616-5600

```

; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-7

```

alignment\_scores:

```

      quality: 592.00      length: 308
      Ratio: 2.779      Gaps: 0
Percent Similarity: 69.156      Percent Identity: 40.260

```

alignment\_block:

US-09-898-586-24 x US-08-748-506-7 ..

Align seg 1/1 to: US-08-748-506-7 from: 1 to: 966

```

3 AspAsnIleThrSerIleThrGluPheLeuLeuGlyPheProValG 19
      ::||| |||::: |||::: |||
28 GAGATAGTTTGTCTGTCAACGCTTTCATTTGCCAAGTTCTCTGAGGT 77
19 yProArgIleGlnMetLeuPheGlyLeuPheSerLeuPheTyrValP 36
      ||| ::||| ||||| |||::: |||:::|
78 CCCTGGAAGATGCTTCCTCTGTCACCTCATCTCTCATGTCTTAG 127
36 heThrLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
      ::||| ||||| ||| ||| |||
128 TATCACTAACAGGAATCTCTCATAGCCCTTGCTATTGTACAGATCCA 177
53 ArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValAs 69
      |||||::: ||||| ||||| ||||| ::::|
178 TCTCTACACACCCCATCTGACTTCCTTGCGCCAACTTCCTCTGGA 227
69 pIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuH 86
      ::|||::: |||::: |||::: |||::: |||:::
228 GATTGGCTAATCTTGTCTGTGTATACCAAGATGCTGCAGAGCCTTGA 277
86 isProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102
      |||||::: |||||::: ||| ||| |||:::
278 GTGAGGCCGAGGATCTTGGGAGGCTGTGCTGCACAGATGTTCTTC 327
103 PheSerThrPheAlaValThrGluCysLeuLeuValValMetSerTy 119
      ||| |||::: ||||| ||||| ||||| ::::|
328 TTCATATCTTGTGTATACTGAGTGCCTGCTATGGCAGCCATGGCCTT 377
119 rAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMe 136
      ::||| |||::: ||||| ||||| ||||| ::::|
378 TGACCGCTATATGCTATATATGTCCTCCACTGCAACCCGAATGA 427
136 hTrpArgValCysIleThrLeuAlaValThrSerTrpThrGlyVal 152
      ::||| ||||| ||||| ||||| ||||| ::::|
428 GTCGTGGGTTATGTGCTATTGGAATTGTCTCATGGGTTGATGGATGC 477
153 LeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPheCysAr 169
      ::|||::: |||::: |||::: |||::: |||:::
478 ATAGTAGTGTGGGACAGACCAATTTTATTTCCTCTGGAACCTCTGTGG 527
169 gProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuLys 186
      ||| ::||| ||||| ||||| ||||| ::::|
528 ACCCTGTGAGATGACCACTTCTCTGTGACCTTCCACCTCTCTGCGAC 577
186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAglAla 202
      ||||| ||||| ||||| ||||| ||||| ::::|
578 TTGCTGTGTGATACATCCCAAAATGAGGCTGCCATCTTTGTGGCAGCA 627
203 IleSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMet 219
      ::|||::: |||::: |||::: |||::: |||:::
628 GTGCTCTGCATATTAGTTCATTTTACTGATCATTTCTTCTATGTGAC 677

```

```
219 sileuencysalaleuGlnleGlnSerArgGluValGlnArgLysA 236
||||| |||::||| ::||| ||| ::::|
678 AATTCTCGTTCAGTGTGATGATGCCCTTACCCTGAGGGCGCCATAAAG 727
236 laphecysthrcyspheserHisleucysValIleGlyLeuPheTyrgly 252
||::| ||||| ||||| |||::| ||||| |||||
728 CTCTCTTACCTGTTCATCTCACCCTACTTGTAGTACACTCTTCTATGGC 777
253 ThrAlaIleIleMetTyrrValGlyProArgTyrglyAsnProLysGluG1 269
:::| |::| |::| |::| |::|
778 TCACATCTGCCACCTATTGTAGGTCCAAGTCTAGCCACTCACAGAGAT 827
269 nLysLysTyrrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
||| ||| |||||:::| ::| |||||
828 GGACAAACTCTTGGCCCTCTTCTATACATCAGTGACATCCATGCTGAATC 877
286 roleuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArg 302
||::| ||| ||||| |||::| ||||| |||::|
878 CCATCATCTACAGCTTAAGAACAGGAAGTAAGGGTGCACCTGAGAGA 927
303 valLeuGlyValGluArgAlaLeu 310
::| |||||:::|
928 ACTCTGGCCCTGAAGAAAGTTCTG 951
```

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-748-506-8

```
seq_documentation_block:
; Sequence 8, Application US/08748506
; Patent No. 6159707
;
; GENERAL INFORMATION:
; APPLICANT: Rommett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
;
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-748-506-8
```

```
alignment_scores:
Quality: 578.00 Length: 308
Ratio: 2.714 Gaps: 0
Percent Similarity: 69.156 Percent Identity: 38.961
```

alignment\_block:
US-09-898-586-24 x US-08-748-506-8 ..

Align seg 1/1 to: US-08-748-506-8 from: 1 to: 966

```
3 AspAsnIlePheHisSerIleThrGluPheLeuLeuGlyPheProValG1 19
::| |::| |::| |::| |::|
28 GAGATAGTTCCTCTGCAAACTTTTGCATTGCGCAAGTCTCTGAGGT 77
19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrrValP 36
||| ::| ||||| |||::| |||::|
78 CCTTGAGATGCTTCCTCCTGTCACCCCTCATCCCTCTCATGTCTTAG 127
36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
::| |::| |::| |::| |::|
128 TATCACTAACAGSAAATGCTCTCATAGCCCTGCTATTGTACCAAGTCCA 177
53 ArgLeuHisAlaProMetTyrrPhePheLeuSerHisLeuAlaValAlas 69
|||||::| ||||| ||||| ||||| |||||
178 TCTCTACACACCCCATGTACTTCTTCTGCCCACACTGTCTCTCTGGA 227
69 PileAlaTyrrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86
::| |::| |::| |::| |::|
228 GATTGCTATACTGCTCTGTCTATACCCACAGATGCTGCAGAGTCTGTGA 277
86 isProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102
||| |::| |::| |::| |::|
278 GTGAGCCCGAGAGATCTTTACGTGGATGTGCCACACAGATGTTTTC 327
103 PheSerThrPheAlaValThrGlyCysLeuLeuValAlaMetSerTy 119
||| |::| |::| |::| |::|
328 TTCATATTCTTGTATTAAGTGAAGTCTGCTATTGGCAGCCATGGCCTT 377
119 rasPleuTyrrValAlaIleCysHisProLeuArgTyrrLeuAlaIleMet 136
::| |::| |::| |::| |::|
378 TGACCGCTATATGGCTATATGTTCCCACTCCACTATGCAACCCGATGA 427
136 hrrTyrrValCysIleThrIleAlaValThrSerTrpThrGlyVal 152
::| |::| |::| |::| |::|
428 GTGCTGAGGTATGTGCCCACTTGCATATTTGATGGATGGATGC 477
153 leuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysAr 169
::| |::| |::| |::| |::|
478 ATAGTAGTCTGGGACAGACCAATTTTATTTCTCTTGAACCTTCTGTGG 527
169 gProGlnLysIleTyrrHisPhePheCysGluIleLeuAlaValLeuLysL 186
||| |::| |::| |::| |::|
528 ACCCTGTGAGATAGACCACTTCTGTGATCTTCCACCTCTCCCGCAC 577
186 euAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAla 202
||| |::| |::| |::| |::|
578 TTGCTGTGCTGATACATCCCAATTTGAGCGTGCATCTTTGTGTAGT 627
203 IleSerGlyLeuValGlyProLeuSerThrIleValAlaSerTyrrMetCy 219
::| |::| |::| |::| |::|
628 GTCTCTGCATATCTAGCCCTTTTGTGCTGATCATTTATTTATGTCAG 677
219 sileuencysalaleuGlnleGlnSerArgGluValGlnArgLysA 236
||||| |||::||| ::||| ||| ::::|
678 AATTCTCGTTCAGTGTGATGATGCCCTTACCCTGAGGGCGCCACAAG 727
236 laphecysthrcyspheserHisleucysValIleGlyLeuPheTyrgly 252
||::| ||||| ||||| |||::| ||||| |||||
728 CCCTTTCAACTGTTCCTCCCACTACTTGTAGTACACTCTTATATGGC 777
253 ThrAlaIleIleMetTyrrValGlyProArgTyrglyAsnProLysGluG1 269
:::| |::| |::| |::| |::|
778 TCAGGATCTGTACCTATTGTAGGCCCTAAGCTAGCCACTCACCAAGAAAT 827
269 nLysLysTyrrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
||| ||| |||||:::| ::| |||||
828 GGACAAACTCTTGGCCCTCTTCTACACAGCAGTGACATCCATGTTGAACC 877
```



```

286 roLeuileCysSerLeuArgAsnSerGluValIysAsnThrLeuIysArg 302
      ||::||| ||||| ||||| ::::||| ||| ::|||::|||
878 CTATCATCTATAGTTTAAAGAACAAAGATGTCAAGCAGCAGCACTGAGAAGA 927
      303 ValLeuGlyValGluArgAlaLeu 310
      ::|||:::||||| |||
928 ATTCTGGCCCTGAAAAAATTCTG 951

seq_name: /cgn2_6/ptodata/2/lna/6A_COMB.seq:US-08-748-506-6

seq_documentation_block:
; Sequence 6, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Meyer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 966 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-6

alignment_scores:
      Quality: 574.00      Length: 308
      Ratio: 2.720      Gaps: 0
Percent Similarity: 68.506      Percent Identity: 39.286

alignment_block:
US-09-898-586-24 x US-08-748-506-6 ..

Align seg 1/1 to: US-08-748-506-6 from: 1 to: 966

3 AspAsnIleThrSerIleThrGluPheLeuLeuGlyPheProValG1 19
      ::||| |||::: ::||| ::| |||
28 GAGAAATAGTTGTCTGTCAAAACGCTTTGCATTGCCAAAGTCTCTGAGGT 77
      19 yProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyValP 36
      ||| ::||| ||||| |||::: |||:::
78 CCCTGGAGAAATGCTTCCTCCTGTTCAACCCFCATTCCTTCATGTTCTTAG 127
      36 heThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSer 52
      ::||| |||||::: |||
128 TATCACTAACAGAGAAATGCTCTCATAGCCCTTGCTGNTGTACCAAGTCCA 177

```

```

53  ATGLeuHHisAlaProMetTyrRhphePheLeuSerHisLeuAlaValAlaS 69
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
178  TCTCTACACACCACCCCATGTACTTCTTCTGGGCCAACCTGTCTCTCTGGA 227
69  pIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuH 86
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228  GATTGGCTATACTTGGCTCTCTGTCAATACCACAGATGCTGCAGAGCCTGTGA 277
86  tSProAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeu 102
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
278  GTGAGGCCACAGAGATCTCTAGGAGGGATGTGCCACACAGATGTTTTTC 327
103  PheSerThrPheAlaValThrGluCysLeuLeuLeuValValMetSerty 119
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328  TTCACATTTTGGCATTAACTAGAGTGGCTGCTATTGGCAGCCATGGCCTT 377
119  rASpLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMetT 136
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
378  TGACCGGTGCATGGGCATATGCTCCCCACTCCACTATGCCAACCCGATGA 427
136  hRTParValCysIleThrLeuAlaValThrSertPrpThrGlyVal 152
    ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
428  GTCGTGAGGATGTGCCCATTTGGCAATTGTTTCATGGGAATGGATGCC 477
153  LeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysAr 169
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478  ATAGTAGGTCTGGGACAGACCAATTNATTTNCTCCTTGAACCTTCTGTGG 527
169  gProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValLeuLysL 186
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528  ACCTGTGAGATAGACCACACTTCTCTGTGACCTTCCACCTCTCCCTGGCAC 577
186  eUAlaCysAlaAspThrHisIleAsnGlnAsnMetValLeuAlaGlyAla 202
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
578  TTGCCCTGTGTGATACATCCCAAAACGAGGCTGCCATCTTGTGGCAGCA 627
203  ILeSerGlyLeuValGlyProLeuSerThrIleValValSerTyrMetCys 219
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
628  ATCCTCTGTATATCTAGTCCATTTTTGGTGATCCTTATTTCTTATGTACAG 677
219  sIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValAlGlnArgLysA 236
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
678  AATTCTCGTTGCAGTGCCTGGTGTATGCTTCACCTGAGGGGGCCCATTAAG 727
236  laPheCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGly 252
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
728  CTCCTTCCACCTGTTCCTCACACCTACTGTAGTCAACACTCTTTTATGGC 777
253  ThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLysGluCl 269
    :|||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
778  TCTGTGTCTTACCCTATTGTGAGGGCCCAAGTCTAGCCACTCACCAGCAT 827
269  nLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnP 286
    :|||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
828  GGACAAACTCTTGGCCCTCTTCTACACAGCAGCATGCATCCATGCTGAACC 877
286  roLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArg 302
    |||:|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
878  CTATCATCTACAGTCTAAGGAACAAGAAGTCAAGCAGCAGCAGTGAAGA 927
303  ValLeuGlyValGluArgAlaLeu 310
    ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
928  ACTCTCGACCTGAATAAATTATG 951

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-085-371-5

seq_documentation_block:
; Sequence 5, Application US/09085371
; Patent No. 6218358
; GENERAL INFORMATION:
; APPLICANT: Firestein, Stuart
; APPLICANT: Zhao, Haiqing

```

```
/ TITLE OF INVENTION: Functional Expression of, and Assay for, Functional Cellular Re
/ TITLE OF INVENTION: vivo
/ FILE REFERENCE: P01511US2 / 09805059
/ CURRENT APPLICATION NUMBER: US/09/085,371
/ CURRENT FILING DATE: 1998-05-19
/ PRIOR APPLICATION NUMBER: US 08/891,243
/ PRIOR FILING DATE: 1997-07-10
/ PRIOR APPLICATION NUMBER: US 60/045,961
/ PRIOR FILING DATE: 1997-05-07
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 5
/ LENGTH: 900
/ TYPE: DNA
/ ORGANISM: Rat
US-09-085-371-5
```

```
alignment_scores:
  Quality: 543.00      Length: 290
  Ratio: 2.675        Gaps: 1
  Percent Similarity: 70.000    Percent Identity: 39.310
```

alignment\_block:  
US-09-898-586-24 x US-09-085-371-5 ..

Align seg 1/1 to: US-09-085-371-5 from: 1 to: 900

```
8  ILeThrGluPheLeuLeuLeuGlyPheProValGlyProArgIleGlnMe 24
   ::::::::::::::::::::
28  GTAGTGAATTGTGTGTGCTGCTCCAGCTGCTGCCACATGCGAGT 77
   ::::::::::::::::::::
24  tLeuLeuPheGlyLeuPheSerLeuPheTyrValPheThrLeuLeuGly 41
   ::::::::::::::::::::
78  ACTACTATTTTCTCTTCCTTCCTGACTATGTGTGTTGTTGACTGAAA 127
   ::::::::::::::::::::
41  snGlyThrIleLeuGlyLeuIleSerLeuAspSerArgLeuHisAlaPro 57
   ::::::::::::::::::::
128  ACATGCTCATCATTTATAGCAATTAGGAACCAACCCCAACCTCCACAACCC 177
   ::::::::::::::::::::
58  MetTyrPhePheLeuSerHisIleuAlaValAlaPheAlaTyrAlaCy 74
   ::::::::::::::::::::
178  ATGTATTTTCTCTGCTAATATATCATTTCTGAGATTGTGATGTAC 227
   ::::::::::::::::::::
74  sAsnThrValProArgMetIleuValAsnLeuLeu.....HisP 87
   ::::::::::::::::::::
228  TGTACGATTCCTAGATGCTGCTGCTCATTTGTTCCAGAGAAC 277
   ::::::::::::::::::::
87  roAlaIysProIleSerPheAlaGlyArgMetMetClnThrPheLeuPhe 103
   ::::::::::::::::::::
278  ATGACAGCTGATCTCTTGTGAGCATGTCATGACACAACCTCTACTTTTTC 327
   ::::::::::::::::::::
104  SerThrPheAlaValThrGluCysLeuLeuIleuValAlaMetSerTyrAs 120
   ::::::::::::::::::::
328  CTGGGCTTGGGTTGCACAGAGTGTCTCTTGTGTGATGGCCATGA 377
   ::::::::::::::::::::
120  pleuTyValAlaIleCysHisProLeuArgTyrIleuAlaIleMetThrT 137
   ::::::::::::::::::::
378  CCGCTATGTGGCTATCTGTCACTCACTCCACTACCCCGTCAATTGTCA 427
   ::::::::::::::::::::
137  rParGValCysIleThrLeuAlaValThrSerTyrThrThrGlyValLeu 153
   ::::::::::::::::::::
428  GCCGGCTATGTGTGAGATGGCAGCTGATCCGCGGTGAGAGTTTGTGT 477
   ::::::::::::::::::::
154  leuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheCysArgPr 170
   ::::::::::::::::::::
478  ATCTCCATGGTTAAAGTTTCTTATTTCTCGCCGTGCTTACTGTGGCCC 527
   ::::::::::::::::::::
170  oGlnIlyIleTyrHisPhePheCysGluIleLeuAlaValLeuIlyLeu 187
   ::::::::::::::::::::
528  CAACACCATCAACCACTTTTCTGTGATGTGTGTTCATTTGCTCAACCTGT 577
   ::::::::::::::::::::
187  lacysAlaAspThrHisIleAsnGluAsnMetValLeuAlaGlyAlaIle 203
```

```
::::::::::::::::::::
578  CATGCACCTGACATGTCCACAGCAGACCTTACAGACTTTGTCTGGCCAT 627
   ::::::::::::::::::::
204  SerGlyLeuValGlyProLeuSerThrIleValAlaSerTyrMetCysIl 220
   ::::::::::::::::::::
628  TTATATTCGTGGGACCGCTCTGTGCACTGGGGCATCTTACATGGCCAT 677
   ::::::::::::::::::::
220  eleuCysAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysAlaP 237
   ::::::::::::::::::::
678  CACAGGTGCTGTGATTCGCATCCCTCAGCTGCTGGCCGCATTAAGCCT 727
   ::::::::::::::::::::
237  heCysThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThr 253
   ::::::::::::::::::::
728  TTTCAACTGTGCTGCCCTCCACCTCAGCTGTGTGATCATCTTATGACACC 777
   ::::::::::::::::::::
254  AlaIleIleMetTyrValGlyProArgTyrGlyAsnProLysGluGlnIly 270
   ::::::::::::::::::::
778  AGTATTTTTCATCTATGCCCAGGCCCTAAGGCACCTCAGCTTTTGACACCAA 827
   ::::::::::::::::::::
270  slySTyrlLeuLeuLeuPheHisSerLeuPheAsnProMetLeuAsnProL 287
   ::::::::::::::::::::
828  CAAGCTGCTCTGTGCTGCTACGCTGTGATGTGTACCGTTGTCAATCCCA 877
   ::::::::::::::::::::
287  euIleCysSerLeuArgAsn 293
   ::::::::::::::::::::
878  TCATCTACTGCTTGGCGAAC 897
```

seq\_name: /cgn2\_6/plodata/2/ina/6A\_COMB.seq:US-08-748-506-9

```
seq_documentation_block:
; Sequence 9, Application US/08748506
; Patent No. 6159707
; GENERAL INFORMATION:
; APPLICANT: Ronnett et al.
; TITLE OF INVENTION: NOVEL SPERM RECEPTORS
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/748,506
; FILING DATE: 08-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,751
; FILING DATE: 09-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; REFERENCE/DOCKET NUMBER: 74940
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5600
; TELEFAX: 312-616-5700
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-748-506-9
```

```
alignment_scores:
  Quality: 452.00      Length: 301
```



```

236 AGATCCCTGCCCTTTCTGCTTTCATTCGCCGAGACATTAGCTTTGAGGCC 285
      :::::  :::  :::::  |||||  :::
96 ArgMetMetGlnThrPheLeuPheSerThrPheAlaValThrGluCysLe 112
      ::  |||  |||:::  :::::  :::
286 TGTCTTACCAGATGTCTCTTATTCATGCCCTTCAGCCATTGAATCCAC 335
      ::  |||  |||:::  :::::  :::
112 uLeuLeuValMetSerTyrAspLeuTyrValAlaIleCysHisProL 129
      ::  |||  |||:::  :::::  :::
336 CATCCTGTGGCCATGTGCGCTTGACCGTTATGTGGCCATCTGCCACCAC 385
      ::  |||  |||:::  :::::  :::
129 euArgTyrLeuAlaIleMetThrTrpArgValCysIleThrLeuAlaVal 145
      ||||:  |||:::  :::::  |||
386 TGGCCATGCTGCAGTGTCAACATACAGTAACAGCCAGATTGGCATC 435
      ||||:  |||:::  :::::  |||
146 ThrSerTrpThrThrGlyValLeuLeuSerLeuIleHisLeuValLeuLe 162
      :::::  :::  |||  |||:::  :::
436 GTGGCTGTGGTCCGGGATCCCTCTTTTTCCTCCACTGCCCTCTGTGAT 485
      :::::  :::  |||  |||:::  :::
162 uLeuProLeuProPheCysArgProGlnLysIleTyrHisPhePheCysG 179
      :  |||  |||||:::  :::  :::  |||  :::
486 CAAGCGGCTGGCCTTCTCTGCCACTCCAATGCTCCTCGCACTCTCATGTG 535
      :  |||  |||||:::  :::  :::  |||  :::
179 LuIleLeuAlaValLeuLysLeuAlaCysAlaAspThrHisIleAsnGlu 195
      |||::  |||||  |||  |||||
536 TCCACCAGATGTATAGAGTGGCTATGCAGACACT..... 573
      |||::  |||||  |||  |||||
196 AsnMetValLeuAlaGlyAlaIleSerGlyLeuValGlyProLeuSerTh 212
      |||  :::::  |||  |||:::  :::
574 .....TTGCCAATGTGTATATGCTGTACTGCCATTCTGTGCT 614
      |||  :::::  |||  |||:::  :::
212 r.....IleValValSerTyrMetCysIleLeuC 222
      :  |||  |||  |||||:::  |||:::
615 CATGGCGGTGGACGTAAATGTTCAATCTCTGTGCTTATTTCTGATATAATAC 664
      :  |||  |||  |||||:::  |||:::
222 ysaAlaIleLeuGlnIleGlnSerArgGluValGlnArgLysAlaPheCys 238
      :::::  |||||:::  |||:::  :::  |||||
665 GAACGGTTCTGCCAACTGCCCTTCCAAAGTCAGAGCGGCCCAAGGCCCTTTGGA 714
      :::::  |||||:::  |||:::  :::  |||||
239 ThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThrAlaIle 255
      |||||  |||||:::  |||:::  |||||
715 ACCTGTGTGCACACATGTGTGTGTACTCGCCTTCTATGTGCCACTTAT 764
      |||||  |||||:::  |||:::  |||||
255 eileMetTyrValGlyProArgTyrGlyAsnProLysGluGlnLysLysT 272
      :  ::  |||  |||::  |||||  :::
765 TGGCCTCTCAGTTGTACACCGCTTTGGAAACAGCCTTCATCCCATTTGTC 814
      :  ::  |||  |||::  |||||  :::
272 yrlLeuLeu.....PheHisSerLeuPheAsnProMetLeuAsnPro 286
      :::::  :::  |||::  |||::  |||||
815 GTGTTGTATGGGTGACATCTACCTGCTGCTGCTCTCTCATCAATCC 864
      :::::  :::  |||::  |||::  |||||
287 LeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArgVa 303
      ::  |||  ::  :::::  :::  |||||
865 ATCATCTATGGTGCACAAACCAACAGATCAGACACA.....CGGGT 905
      ::  |||  ::  :::::  :::  |||||
303 lLeuGlyVal 306
      |||:::  :::
906 CCTGGCTATG 915

```

seq\_name: /cgn2\_6/prodata/2/lna/6A\_COMB.seq:US-08-988-876-2

seq\_documentation\_block:

```

; Sequence 2, Application US/08988876
; Patent No. 6063596
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTORS ASSOCIATED
; WITH IMMUNE RESPONSE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,876
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0441 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1828 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT01
; CLONE: 364702
; US-08-988-876-2

```

#### alignment\_scores:

Quality:	382.50	Length:	307
Ratio:	1.922	Gaps:	8
Percent Similarity:	64.821	Percent Identity:	37.785

#### alignment\_block:

US-09-898-586-24 x US-08-988-876-2 ..

Align seg 1/1 to: US-08-988-876-2 from: 1 to: 1828

```

4 AsnIleThrSerIleThrGluPheLeuLeuGlyPheProValGlyPr 20
  |||::  |||::  :::::  |||||  |||
793 AATCTAACAGATGCTCTATATTCCTCTCCTAGAGTCTCAGGGGATCC 842
  20 oArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheTyrValPheT 37
  ||:::  |||  ::  |||  |||||  :::
843 AGAACTGCAGCCAGCTGCTGCTGCTGCTGCTGCTGCCATGTCAGCTGCCAC 892
  37 hrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeuAspSerArg 53
  ||:::  |||||  |||::  |||  |||||  |||||
893 CGGTGCTGGGAACCTGCTCATCATCTGGCCATCAGACCTGACTGCCAC 942
  54 LeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaValAlaAspI 70
  |||||  ::  |||||  |||||  |||||  ::  |||||
943 CTCACACACCCCATGTACTTCTTCTCTTCCATCAACCTGTGCTGCTGACAT 992
  70 eAlaIlyAlaCysAsnThrValProArgMetLeuValAsnLeuLeuHisP 87
  ||:::  :::  ::  |||||  |||::  |||  |||||
993 CGGTTTCACCTCCACACCGGTCCCAAGATGATGTGTGACATCCAGTCTC 1042
  87 roAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPheLeuPhe 103
  :::::  |||||  ::  |||||  :::  |||  |||||
1043 ACAGCAGAGTCACTCTCTATGCAAGCGTGCCTGACTCAGATGCTCTCTT 1092
  104 SerThrPhe..AlaValThrGlu.CysLeuLeuLeuValValMetSert 119

```

Fri Aug 23 16:18:01 2002

us-09-898-586-24.inl

Page 11

1093	GCCATTTTGTGGAGCATGGAAGAGACATGCTCCATGAGTGTGATGGCT	1142
119	yrAspLeuTyValAlaIleCysHisProLeuArgTyLeuAlaIleMet	135
1143	ATGACTGGTGTGTAGCCATCTGTACACCGCTATATCATTC, ACCATCATG	1191
136	ThrTrpArgValCysIleThrLeuAlaValThrSer... TrpThrThrGl	151
1192	AACCCGCTGTCTGTGGCTTCTCTGACTTTTGTGTCTTTTCTTCTCAG	1241
151	ValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPheC	168
1242	TCTTTTAGACTGCCAGCTGCACAACTTGATTTGCCCTTACAAGTGACCTGCT	1291
168	ysArgProGlnLysIleTyHisPhePheCysGlnIleLeuAlaValLeu	184
1292	TCAAGGATGTGGAATTCCTTAATTTCTTCTGTGACCTTCTCAACTCTCC	1341
185	LysLeuAlaCysAlaAspThrHis... IleAsnGluAsnMetValLeuAl	200
1342	CATCTTGCACTGTGTGACACCTTCACCATTAACATATATCATGTATTTCCTCC	1391
200	agLysAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerT	217
1392	TGCTGCCATATTTGGTTTCTT... CCCATCTCAGGACCTTTCTCTTTA	1438
217	yrMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln	233
1439	CTGTAAATAATCTTCTCCATCTGTGAGGGTTTCATCATCAGGTGGGAG	1488
234	ArgLysAlaPheCysThrCysPheSerHisLeuCysValIleGly, LeuP	250
1489	TATTAACC, TTCTCCACCTGTGGGTCTCACCTG, ..... TCAGTTGT	1528
250	heTyArgLyThrAlaIleIleMetTyValGlyProArgTyGlyAsnPro	266
1529	TTGCTGGAGG, ..... TACCTCGTTACAGTGTGTCACTTCTCC	1566
267	LysGlnGlnLysLysTyLeuLeuLeuPheHisSerLeuPheAsnProMe	283
1567	CCGAGAAAGAGTGCAGTGGCTCAGTATGTACACGGTGGTCACCCCAT	1616
283	tLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrL	300
1617	GCTGAACCCCTTCATGTACAGCCTGAGAAACAGGGATATGAAGAAGTCTC	1666
300	eulYsArgValLeuGly 305	
1667	TGCGGCGCGCGCACGGC 1683	

```
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-465-980-1
```

```

seq_documentation_block:
; Sequence 1, Application US/08465980
; Patent No. 5756309
;
; GENERAL INFORMATION:
;
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
;
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRA170
;
; NUMBER OF SEQUENCES: 8
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
;
; ZIP: 07068
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;

```

```

1 COMPUTER: IBM PC compatible
2 OPERATING SYSTEM: PC-DOS/MS-DOS
3 SOFTWARE: PatentIn Release #1.0, Version #1.30
4 CURRENT APPLICATION DATA:
5 APPLICATION NUMBER: US/08/465,980
6 FILING DATE: 06-JUN-1995
7 CLASSIFICATION: 536
8 ATTORNEY/AGENT INFORMATION:
9 NAME: Ferraro, Gregory D.
10 REGISTRATION NUMBER: 36,134
11 REFERENCE/DOCKET NUMBER: 325800-446
12 TELECOMMUNICATION INFORMATION:
13 TELEPHONE: 201-994-1700
14 TELEFAX: 201-994-1744
15 INFORMATION FOR SEQ ID NO: 1:
16 SEQUENCE CHARACTERISTICS:
17 LENGTH: 1474 base pairs
18 TYPE: nucleic acid
19 STRANDEDNESS: single
20 TOPOLOGY: linear
21 MOLECULE TYPE: cDNA
22 FEATURE:
23 NAME/KEY: CDS
24 LOCATION: 274..1233
25
26 US-08-465-980-1

```

alignment_scores:		
Quality:	369.00	Length: 320
Ratio:	1.883	Gaps: 8
Percent Similarity:	61.250	Percent Identity: 30.625

```
alignment_block:
US-09-8998-586-24 x US-08-465-980-1 .
```

Align seg 1/1 to: US-08-465-980-1 from: 1 to: 1474

```

4 AsnIleThrSerIleThrgluPheLeuLeuGlyPhePro..... 17
   |||::||| ||| ::|||::|||::|||
286 AACTTCACACATGCCACC..TGTCGCTTATGTGTATCCACGATTAGA 332
18 .....ValGlyProArgIleGlnMetLeuLeuPheGlyL 29
   ||||| ||| |
333 GAAAGCCCATTTCTGGGTTGGC.....TTCCCCC 361
29 eupHeserLeuPheTyrrValPheThrLeuLeuGlyAsnGlyThrIleLeu 45
   ||::|||:: ||||| :::: ||||| ::::
362 TCCTTTCCATG..TATGTAGTGGCAATGTGTGGAACATGCATCGTGGTC 408
46 GlyLeuIleSerLeuAspSerArgLeuHisAlaProMetTyrrPhePhe 62
   :::: :: ||||| |||||::|||
409 TTCATCGTAAGACGGAACGACGCGCTGCACGCTCCGATGTACCTCTTCT 458
62 userHisLeuAlaValAlaSplIleAlaTyrrAlaCysAsnThrValPro 79
   | |||||::: |||:: ||| :::: |||::
459 CTGCATGCTTGCAGCATTGACCTGGGCTTATCCACATCCACCATGGCTTA 508
79 rgMetLeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGly 95
   :::: ||| :::: |||||:::
509 AGATCCCTTGCCCCCTTCTGTTGATTGATCCCGAGAGATTAGCATTTGAGGC 558
96 ArgMetMetGlnThrPheLeuPheSerThrPheAlaValThrgluCysIle 112
   :: ||| |||::: ::::: |||
559 TGTCCTTACCACAGATCTTCTTATTCATGACCTCTCAGCCATTGAAATCCAC 608
112 uLeuLeuValValMetSerTyrrAspLeuTyrrValAlaIleCysHisPro 129
   ::|||::: |||:: ||| ||||| ||||| |||||
609 CATCCCTGCTGGCCCATGGCCCTTTGACCGTTATGTGGCCATCTGCCACCCAC 658
129 euArgTyrrLeuAlaIleMetThrTrpPargValCysIleThrLeuAlaVal 145
   |||||:: |||:: ||| :::::
659 TGGGCCATGCTGCGAGTCTCAACACATACAGTAACAGCCACGATTGGCATC 708

```



```
146 ThrSerTrpThrThrGlyValLeuLeuSerLeuLeuHisLeuValLeu 162
      ::::: ::::: ||| |||::: ::::: ::::: |||:::
709 GTGGCTGTGGTCCGGGATCCCTCTTTTTCCTCCACTGCCTCTGCTGAT 758
      : ||| |||::: ::::: ::::: ||| |||:::
162 uLeuProLeuProPheCysArgProGlnLysIleTyrHisPhePheCys 179
      : ||| |||::: ::::: ::::: ||| |||:::
759 CAAGCGGCTGGCCCTCTCCCACTCCCAATGCTCTCCGCACTCCATGTTG 808
      : ||| |||::: ::::: ::::: ||| |||:::
179 LuLeuAlaValLeuLysLeuAlaCysAlaAspThrHisIleAsnGlu 195
      |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
809 TCCACCAGAGATCTAATGAAGTTGGCCCTATGCAGACACT..... 846
196 AsnMetValLeuAlaGlyAlaIleSerGlyLeuValGlyProLeuSerTh 212
      ||| ::::: ||| |||::: ||| |||::: ||| :::::
847 .....TTGCCCAATGTGTATATGTGCTTACTGCCATTCTGCTGCT 887
212 T..... |||ValSerTyrMetCysIleLeu 222
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
888 CATGGCGGTGACGTAATGTTCATCTCCTGTGCTATTTTCTGATATAATAC 937
222 ysaIaIleLeuGlnIleGlnSerArgGluValGlnArgLysAlaPheCys 238
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
938 GAACGGTCTGCAACTGCCCTCCCAAGTCAGAGCGGCGCAAGGCCTTTGA 987
239 ThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThrAlaI 255
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
988 ACCTGTGTGTACACATTTGGTGTGCTACTCGCCTTCTATGTGCACCTAT 1037
255 eileMetYrValGlyProArgTyrGlyAsnProLysGluGlnLysIle 272
      | ::::: ||| |||::: ||| ||| ||| ||| ||| ||| |||
1038 TGGCCTCTCAGTTGTACACCGCTTTGAAACAGCCCTTCATCCCATTTGTC 1087
272 yrLeuLeuLeu.....PheHisSerLeuPheAsnProMetLeuAsnPro 286
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1088 GTGTGTGTATGGGTGACATCTACCTGCTGCTGCTCTCTGTCATCAATCCC 1137
287 LeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeuLysArgVa 303
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1138 ATCATCTATGTGCCCAAAACCAACAGATCAGACACA.....CGGGT 1178
303 lleuGlyVal 306
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1179 GCTGGCTATG 1188

seq_name: /cgn2_6/ptodata/2/lna/5B_COMB.seq:US-09-053-303-1

seq_documentation_block:
; Sequence 1, Application US/09053303
; Patent No. 5948890
;
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAJ70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESS: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,303
; FILING DATE:
; CLASSIFICATION:
```

```
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/465,980
;; FILING DATE: 06-JUN-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Ferraro, Gregory D.
;; REGISTRATION NUMBER: 36,134
;; REFERENCE/DOCKET NUMBER: 325800-446
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 201-994-1700
;; TELEFAX: 201-994-1744
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1474 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: cDNA
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: 274..1233
;;
;; US-09-053-303-1

alignment_scores:
      Quality: 369.00      Length: 320
      Ratio: 1.883      Gaps: 8
Percent Similarity: 61.250      Percent Identity: 30.625

alignment_block:
US-09-898-586-24 x US-09-053-303-1 ..

Align seg 1/1 to: US-09-053-303-1 from: 1 to: 1474

4 AsnIleThrSerIleThrGluPheLeuLeuGlyPhePro..... 17
      |||::: ||| ||| ::::: ||| ||| ||| ||| ||| ||| |||
286 AACTTCACACATGCCACC...TGTGTCTTATTGTATCCACAGATTAGA 332
18 .....ValGlyProArgIleGlnMetLeuLeuPheGlyL 29
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 GAAAGCCCATTTCTGGGTTGGC.....TTCCCCC 361
29 euhPheSerLeuPheTyrValPheThrLeuLeuGlyAsnGlyThrIleLeu 45
      |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TCCTTTTCCATG...TATGTAGTGGCAATGTGTGAAACTGCAATCGTGTGTC 408
46 GlyLeuIleSerLeuAspSerArgLeuHisAlaProMetTyrPhePheLe 62
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
409 TTCAATCGTAAGACGGAACGACGCTGCACGCTCCGATGTACCTCTTCT 458
62 uSerHisLeuAlaValAlaAspIleAlaTyrAlaCysAsnThrValProA 79
      | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
459 CTGCATGCTTGACGCACTGACCTGCTTATCCACATCCACATGCGCTA 508
79 rGMetLeuValAsnLeuLeuHisProAlaLysProIleSerPheAlaGly 95
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
509 AGATCCCTTGCCCTTTCTGTGCTTTGATGCCGAGAGATTAGCATTTGAGGCC 558
96 ArgMetMetGlnThrPheLeuPheSerThrPheAlaValThrGluCysLe 112
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
559 TGTCTTACCACAGATGTTCTTTATTATTCATGCCCTCTCAGCCATTGAATCCAC 608
112 uLeuLeuValValMetSerTyrAspLeuTyrValAlaIleCysHisProL 129
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
609 CATCTGCTGGCCATGGCCCTTTGACCCGTTATGTGGCCATCTGCCACCCAC 658
129 euArgTyrLeuAlaIleMetThrTrpArgValCysIleThrLeuAlaVal 145
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
659 TGGCGCATGTGTGACGTCTCAACAATACAGTAACAGCCCAAGTTGGCATC 708
146 ThrSerTrpThrThrGlyValLeuLeuSerLeuIleHisLeuValLeuLe 162
      : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
709 GTGGCTGTGTCCGGGATCCCTCTTTTTCCTCCACTGCCTCTGCTGAT 758
```

```

162 uleuProleupropheCysargProGlnLysIleTyrHisPhePheCysG 179
: ||| |||||::: :::: ||| ::|||
759 CAAGCGGCTGGCTTCTGCCACTGCCAATGTCCTTCGCCACTCCTATTGTG 808
179 luileuAlaValleuLysleuAlaCysAlaAspThrHisIleAsnGlu 195
|||:::||||||| |||||
809 TCCACCAGAGATGTAATGAAGTGGCTTATGCAGACACT..... 846
196 AsnMetValleuAlaGlyAlaIleSerGlyLeuValGlyProleuSerTh 212
||| ::::: |||||::: ||| ::
847 .....TTGCCAATGTGTATATGGTCTTACTGCCATTCTGCTGTG 887
212 r.....IleValIleSerTyrMetCysIleLeu 222
: ||| ::|||::: |||:::
888 CATGGCGGTGACGTAATGTTCACTCTCTGCTTCTTCTGATTAATAC 937
222 ysaAlaIleuGlnIleGlnSerArgGluValGlnArgLysAlaPheCys 238
:::|||||::: |||::: :: |||||
938 GAACGGTCTTGCACACTGCCCTTCCAGTCAGAGCGGCCCAAGCCTTTGGA 987
239 ThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThrAlaI 255
||||| |||||::: |||::: |||||
988 ACCTGTGTGTACACATTGGTGTGTACTCGCCTTCTATGTGCCACTTAT 1037
255 eileMetTyrValGlyProArgTyrGlyAsnProLysGluGlnLysLysT 272
|::: ||| |||::|||:::
1038 TGGCCTCTCAGTTGTACACCGCTTTGGAAACAGCCTTTCATCCCATTTGTC 1087
272 yrleuLeu.....PheHisSerleuPheAsnPrometleuAsnPro 286
::: |||::: |||::: |||::: |||||
1088 GTGTGTGTCATGGGTGACATCTACCTGCTGCTGCCCTCTGTCATCAATCCC 1137
287 leuIleCysSerleuArgAsnSerGluValLysAsnThrleuLysArgVa 303
::||| ::::: |||::: |||||
1138 ATCATCTATGTGGCCAAACCAACAGATCAGAC.....CGGGT 1178
303 lleuGlyVal 306
||||:::
1179 GCTGGCTATG 1188

```

seq\_name: /cgn2\_6/ptodata/2/lna/PCTUS\_COMB.seq:PCT-US95-07093-1

```

seq_documentation_block:
; Sequence 1, Application PC/TUS9507093
; GENERAL INFORMATION:
; APPLICANT: Soppet, Daniel R.
; APPLICANT: Li, Yi
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: HUMAN G-PROTEIN RECEPTOR HPRAU70
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STUART & OLSTEIN
; STREET: 6 Becker Farm Road
; CITY: Roseland
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/07093
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/465,980
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferraro, Gregory D.

```

```

; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-446
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 274..1233
; PCT-US95-07093-1

```

```

alignment_scores:
Quality: 369.00 Length: 320
Ratio: 1.883 Gaps: 8
Percent Similarity: 61.250 Percent Identity: 30.625

```

alignment\_block:

US-09-898-586-24 x PCT-US95-07093-1 ..

Align seg 1/1 to: PCT-US95-07093-1 from: 1 to: 1474

```

4 AsnIleThrSerIleThrGluPheLeuLeuGlyPhePro..... 17
||||:|||| ||| ::|||::|||
286 AACTTCACACATGCGCAC..TGTTGCTTATTTGTTATCCAGATTAGA 332
18 .....ValGlyProArgIleGlnMetleuLeuPheGlyL 29
|||||
333 GAAGCCCATTTCTGGCTTGGC.....TTCCCCC 361
29 eurPheSerleuPheTyrValPheThrleuLeuGlyAsnGlyThrIleleu 45
||::|||::: ||||| ::::: ||||| :::::
362 TCCCTTTCATG..TATGTAGTGGCAATGTGTGAACCTGCATCGTGGTC 408
46 GlyleuIleSerleuAspSerArgleuHisAlaPrometTyrPhePhe 62
::: |||::: ||| ||||| |||||::: |||||
409 TTCATCGTAAAGAGCGAAGCAGCGCTGCACGCTCCGATGTACCTTTCT 458
62 userHisleuAlaValAlaAspIleAlaTyrAlaCysAsnThrValProA 79
| |||||::: |||::: |||::: |||:::
459 CTGCATGCTTGCAAGCATTGACCTGGCTTATCCACATCCACCATGGCTTA 508
79 rgMetleuValAsnleuLeuHisProAlaLysProIleSerPheAlaGly 95
::: |||::: ::: ::::: |||||:::
509 AGATCCCTTGCCCTTTCTGTTGTGATTCGCGAGAGATAGCATTGAGGCC 558
96 ArgMetMetGlnThrPheleuPheSerThrPheAlaValThrGluCysIle 112
::: ||| |||::: |||::: |||::: |||
559 TGCTCTTACCCAGATGTTCTTTATTCATGCCCCCTCAGCCATTGAATCCAC 608
112 uleuLeuValValMetSerTyrAspLeuTyrValAlaIleCysHisProL 129
::: |||::: |||::: ||| ||||| ||||| ||||| |||||
609 CATCCCTGCTGGCCATGGCCCTTTGACCGGTTATGTGGCCATCTGCCACCCAC 658
129 euArgTyrleuAlaIleMetThrTyrPargValCysIleThrleuAlaVal 145
|||||::: |||::: ||| |||
659 TGGCCCATGTGCTGAGTGTCAACAATACAGTAAACAGCCAGATTGGCACC 708
146 ThrSerTyrThrThrGlyValleuLeuSerleuIleHisleuValleu 162
::: |||::: ||| |||::: |||::: |||:::
709 GTGGCTGTGTCCGGGATCCCTCTTTTCTTCCCACTGCTCTGTGTGAT 758
162 uleuProleupropheCysargProGlnLysIleTyrHisPhePheCysG 179
: ||| |||||::: :::: ||| ::|||
759 CAAGCGGCTGGCTTCTGCCACTGCCAATGTCCTCTGCGCACCTCCTATTGTG 808
179 luileuAlaValleuLysleuAlaCysAlaAspThrHisIleAsnGlu 195

```

Fri Aug 23 16:18:01 2002

us-09-898-586-24.rn

Page 14

809	TCCACGAGATGTAATGAGAGTTGGCTATGCAGACACT		846
196	AsnMetValLeuAlaGlyAlaIleSerGlyLeuValGlyProLeuSerTh		212
847	.....TTGCCCAATGTGGTATATGTGCTTACTGCCATTCTGTGCTGT		887
212	r.....IleValValSerTyrMetCysIleLeu		222
888	CATGGGGCGTGAACGTAATGTTCATCTTCCTGTGCTTATTTCTGATATATC		937
222	ysAlaIleLeuGlnIleGlnSerArgIleValGlnArgIlyAlaPheCys		238
938	GAACGGTTCTGCAACTGCCTTCCAAAGTCAGAGCGGGCCAAAGCCTTTGGA		987
239	ThrCysPheSerHisLeuCysValIleGlyLeuPheTyrGlyThralaI		255
988	ACCTGTGTGTACACATTTGGTGTGTACTCGCCCTTCTATGTGCCACTTAT		1037
255	eIleMetTyrValGlyProArgTyrGlyAsnProIlySGlnIlyLysIst		272
1038	TGGCCTCTCAGTTGTATCACACCGCTTGGAAACAGCCCTTTCATCCCAATTGTGC		1087
272	yrLeuLeu.....PheHisSerLeuPheAsnProMetLeuAsnPro		286
1088	GTTGTGTCAATGCGCTGACATCTTACCCTGCTGCTGCCCTTCCTGTCATCAATCC		1137
287	LeuIleCysSerLeuArgAsnSerGluValIlyAsnThrLeuLysArgVa		303
1138	ATCATCTATAGGTGCCCAAAACCAACAGATCAGAACA.....CGGGT		1178
303	IleuGlyVal 306		
1179	GCTGGCTATG 1188		

seq\_name: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:US-08-706-281A-17

seq\_documentation\_block:  
: Sequence 17, Application US/08706281A  
: Patent No. 6100048  
: GENERAL INFORMATION:  
: APPLICANT: Cone, Roger D  
: APPLICANT: Fan, Wei  
: APPLICANT: Boston, Bruce A  
: APPLICANT: Kesteron, Robert A  
: APPLICANT: Lu, Dongxi  
: APPLICANT: Chen, Wenbiao  
: TITLE OF INVENTION: Methods and Reagents for Discovering and  
: TITLE OF INVENTION: Using Mammalian Melanocortin Receptor Agonists and Antagonists  
: TITLE OF INVENTION: To Modulate Feeding Behavior in Animals  
: NUMBER OF SEQUENCES: 19  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff  
: STREET: 300 South Wacker Drive  
: CITY: Chicago  
: STATE: IL  
: COUNTRY: USA  
: ZIP: 60606  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/706,281A  
: FILING DATE: 04-SEP-1996  
: CLASSIFICATION: 435  
: ATTORNEY/AGENT INFORMATION:  
: NAME: No. 6100048nan, Kevin E  
: REGISTRATION NUMBER: 35,303  
: REFERENCE/DOCKET NUMBER: 96,886  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 312-913-0001

```

/ TELEFAX: 312-913-0002
/
/ TELEX:
/
/ INFORMATION FOR SEQ ID NO: 17:
/
/ SEQUENCE CHARACTERISTICS:
/
/ LENGTH: 978 base pairs
/
/ TYPE: nucleic acid
/
/ STRANDEDNESS: single
/
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: DNA (genomic)
/
/ FEATURE:
/
/ NAME/KEY: CDS
/
/ LOCATION: 1..975
/
US-08-706-281A-17

```

```
alignment_scores:
  Quality: 185.50      Length: 305
  Ratio: 1.124         Gaps: 13
Percent Similarity: 54.098  Percent Identity: 24.918
```

alignment\_block:

Align seg 1/1 to: US-08-706-281A-17 from: 1 to: 978

25 LeuleuPheGlyLeuPheSerLeuPheTyrValPheThrLeuGlyAs 41  
||| ::::| | | | | ||| | | | |  
130 CTGACCCCTGGGTCTCGTCAGC.....CTCTTAGAGAA 161

41 nGlyThrIleLeuGlyLeuIleSerLeuAspSerArgLeuHisAlaProM 58  
| : ::::| | | | | ||| : : : : : | | | | |  
162 CATCCCTGCATTTGGGCCCATGTAAAGCAACAAAAACCCTGCACTCACCCA 211

58 eTyrPhePheLeuSerHisLeuAlaValAlaSp.....Ile 70  
| | | | | | | | | | : : | | | | | | | | | |  
212 TGTACTTCTTTGTGGGACGCTTAGCCGTGGCCGACATGCTGCTGAGCATG 261

71 AlaTyrAlaCysAsnThrValProArgMetLeuValAsnLeuLeuHispr 87  
::: | | | : : | | | | | | | | | | | | | |  
262 TCCAATGCTTGAGAGACTGTCCACCATATACTTGCTAATAATAAACACCT 311

87 oAlaLysProIleSerPheAlaGlyArgMetMetGlnThrPhe....L 102  
:::  
312 GGTGAATAGCCGACACACTTTGTGCGACACATCGACAACGNTTGCAGCTCCA 361

102 euPheSerThrPheAlaValThrGlucylsLeuLeuValAlaMetSer 118  
::::: : : | | | | : : : : : : : : : : : : : : : :  
362 TGATCTGCATCTCTGTGGTGGCCCTCGATGTCAGTTTGGCTGGCCATTGGC 411

119 TyrAspleuTyrValAlaIleCysHisProLeuArgTyrLeuAlaIleMe 135  
||| | | | : : : : : : : : : : | | | | | | | | | |  
412 GTGATAGGTATACATCACCATCTTCTATGCCTTGGCGGTACCAACCATCAT 461

135 tThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThrGlyV 152  
||| | | | : : : : : : : : : : | | | | | | | | | |  
462 GACCGGAGGCGCTCGGGGGTGATCATGCGCTGCATTTGCAGCCTTCTGCA 511

152 alLeuLeuSerLeuIleHisLeuVal..... 160  
:::  
512 TAAGCTCGGGCAATGTTTTTCATCATCTACTATGAGTCCAGTATGTGATC 561

161 .....LeuLeuLeuProLeuProPheCysArgProGlnIy 172  
:  
562 ATTGCGCTCATCTGCAGATGTTCTTACACCATGCTGTCTTGCATGGTGTCT 611

172 sIleTyrHisPhePheCysGluIleLeuAla.....ValLeuIysl 186  
||| : : | | | : : | | | | | | | | | | | | | |  
612 GTATATTACACATGTTCC.....CTCCTGGCCCGGAACCATGTCAAGCGCA 655

186 euAlaCysAlaAspThrHis.....IleasnGluAsnMetValLeuAla 200  
::| | | :  
656 TAGCAGCTTCCCCCAGATAACAACCTCCGTGAGGCCAAGAAGCCAGCATGAAG 705



11-11-11

1

2

3





```

84 euleuHisProAlaIysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
680 TCCTGCATCCAGCAAGCCCATCTCTTGGCTGGTTCATGACGAGACC 631
101 PheLeuPheSerThrPheAlaValThrGlyCysLeuLeuValAlaMe 117
|||||
630 TTTCCTGTGTTGAGTTTGGACACAGCGAATGTCTCTGCTGGTGTGAT 581
117 tSerTyrrAspLeuTyrrValAlaIleCysHisProLeuArgTyrrLeuAlaI 134
|||||
580 GTCTACGATCGTTACGTGGCCATCTGCCACCTCTCCGATACCTCCGTC 531
134 leMetThrTrpArgValCysIleThrIleuAlaValThrSerTrpThrThr 150
|||||
530 TCATGACCTGGAGAGTGTGCATCACCCTGGCCGTCACTTCTGTGAGGTGT 481
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167
|||||
480 GGCTCCCTCTCTGGCTGTGGCCCATGTGGTTCATCTTAAGACTGCCCTT 431
167 eCysArgProGlnIlySileTyrrHisPhePheCysGluIleLeuAlaVal 184
|||||
430 CTCTGGGCCCTCATGAATCAACCACTTCTTGTGTGAATCCTGTCTGTCC 381
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||||
380 TCAGGCTGGCCTGTGCTGACACCTGGCTCAACCAAGTGTTCATCTTTGCA 331
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
|||||
330 GCCTGGCTGTCTCTCTGTGGGGGCCACCCAGCCTGTGTGTCTCTTA 281
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
|||||
280 CTGCAATCTCTGGCGGCCATCTGAGGATCCAGTCTGGGGAGGGGCCGA 231
234 rGlySAlaPheCysThrCysPheSerHisIleCysValIleGlyLeuPhe 250
|||||
230 GAAAGGCTTCTCCACCTGCTCCTCCACCTGTGGGTGTGGAGCTCTTC 181
251 TyrGlyThrAlaIleIleMetTyrrValGlyProArgTyrrGlyAsnProLy 267
|||||
180 TTTGGCAGTGCATCATCATGATGCCCCCAAGTCCCGCATCTCTGA 131
267 sGluGlnIlySlyTyrrLeuLeuPheHisSerLeuPheAsnProMetL 284
|||||
130 GGAGCAGCAAAAGTCTTTTCTTATTTTACAGTTTTCACACCCACAC 81
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValIlyAsnThrLeu 300
|||||
80 TTAACCCCTGATTTACAGCCTGAGGAACGAGAGTCAAGGTGCCCTG 31
301 LysArgValLeuGlyValGlu 307
:::|||||
30 AGGAGAGCACTGGGCAAGGAA 10
```

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAI23185

seq\_documentation\_block:

ID AAI23185 standard; DNA; 933 BP.

XX AAI23185;

DT 12-OCT-2001 (first entry)

DE Probe #13118 for gene expression analysis in human cervical cell sample.

XX KW Probe; human: microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer; ss.

OS Homo sapiens.

XX WO200157278-A2.

```

XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00670.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPL; 2001-488901/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
PS Claim 25; SEQ ID NO 13118; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENP). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging
CC of diseases of the cervix, notably cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ
XX Sequence 933 BP; 243 A; 220 C; 303 G; 167 T; 0 other;

alignment_scores:
Quality: 1126.00 Length: 307
Ratio: 4.155 Gaps: 0
Percent Similarity: 88.274 Percent Identity: 71.336

alignment_block:
US-09-898-586-24 x AAI23185/rev ..

Align seg 1/1 to reverse of: AAI23185 from: 1 to: 933

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||
930 ATGGGGGAATCAGACAATGTCACAGAGTTCCTCTACTGGGATTTCT 881
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
:::|||||
880 CCTGGGCCCAAGATTCAGATGCTCCTTGGGCTCTTCCCTGTTCT 831
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
|||||
830 ATATCTTCACTGCTGTGGGGAACGGGGCCATCTGGGGCTCATCTCACTG 781
51 AspSerArgLeuHisAlaProMetTyrrPhePheLeuSerHisLeuAlaVal 67
|||||
780 GACTCCAGACTCCACACCCCATGTACTTCTCTCTCCACACACCTGGCTGT 731
67 lValAspIleAlaTyrrAlaCysAsnThrValProArgMetLeuValAsnL 84
|||||
730 CGTCGACATCGCTTACACCGGCAACAGGTTGCCCATGTGCGGAAC 681
84 euleuHisProAlaIysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
680 TCCTGCATCCAGCCAAAGCCATCTCTTGTGTGATGATGACGACAGCC 631
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValAlaMe 117
```

|||||  
780 GACTCCAGACTCCACACCCCATGTACTCTCTCTCCACACCTGGCTGT 731  
67 lValaspIleAlaTYrAlaCysAsnThrValProArgMetLeuValAsnL 84  
|||||  
730 CGTCGACATCGCCTACACCCGCAACACGGTGGCCAGATGCTGGCAACC 681  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
680 TCCTGCATCCAGCCAGCCCATCTCCTTGCTGGTTGCATGACGACAGACC 631  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
630 TTTCTCTGTGTAGATTGTGGACACAGCGAAATGTCCTGCTGGTGTGAT 581  
117 tSerTYrAspLeuTYrValAlaIleCysHisProLeuArgTYrLeuAlaI 134  
|||||  
580 GTCTTAGCGATCGTTACGTGGCCATCTGCCACCCCTCCGATACCTCCGTC 531  
134 lMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
|||||  
530 TCATGACCTGGAGAGTCTGCATCACCCTGGCCGTCACCTCCTGGACGTGT 481  
151 GlyValIleuLeuSerIleuHisLeuValIleuLeuProLeuProph 167  
|||||  
480 GGCTCCCTCCTGGCTGTGGCCATGTGCTCTCATCTAAGACTGCCCTT 431  
167 eCysArgProGlnLysIleTYrHisPhePheCysGluIleLeuAlaValL 184  
|||||  
430 CTCTGGCCCTCATGAATCAACACACTCTCTGTGAATCCGTCTGTCC 381  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValIleuAla 200  
|||||  
380 TCAGCTGGCTGTGCTGACACCTGGCTCAACAGGTGTGCATCTTTGCA 331  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTY 217  
|||||  
330 GCCTGGCTGTCTTCTCTGGTGGGGCCACCCAGCTGTGCTGTCTCTTA 281  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
|||||  
280 CTGCGACATCCTGGCGGCCATCCTGAGGATCCAGTCTGGGAGGGCCGCA 231  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
230 GAAAGGCTCTTCACACCTCTCTCCACCTCTGCGTGGGACTCTTTC 181  
251 TYrGlyThrAlaIleIleMetTYrValGlyProArgTYrGlyAsnProLy 267  
|||||  
180 TTTGGCAGTGCATCAATGATGATGAGCCGCCCAAGTCCCGCATCTCTGA 131  
267 sGluGlnLysLysTYrIleuLeuLeuPheHisSerLeuPheAsnPrometL 284  
|||||  
130 GGAGCAGCAAAAGGTCTTTTCTATTTTACAGTTTTCACCCCAACAC 81  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||||  
80 TTAACCCCTGATTACAGCTGAGGAACGAGAGGTCAAGGTGCCCTG 31  
301 LysArgValLeuGlyValGlu 307  
|||||  
30 AGGAGAGCACTGGCGAAGAA 10

seq\_name: /sids1/gcgdata/hold-geneseq/genesegn-emb1/NA2001A.DAT:ABA68294

seq\_documentation\_block:

ID ABA68294 standard; DNA; 933 BP.

XX ABA68294;

AC

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #16599.

XX Human: foetal liver; gene expression; single exon nucleic acid probe; ss.  
KW Homo sapiens.  
XX WO200157277-A2.  
OS 09-AUG-2001.  
PD 30-JAN-2001; 2001WO-US00669.  
PF 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-483447/52.  
DR Human genome-derived single exon nucleic acid probes useful for  
XX analyzing gene expression in human fetal liver -  
PT Claim 4; SEQ ID NO 16599; 639pp + sequence listing; English.  
PS The invention relates to a single exon nucleic acid probe for  
XX measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a single exon nucleic acid  
CC probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX Sequence 933 BP; 243 A; 220 C; 303 G; 167 T; 0 other;  
SQ  
alignment\_scores:  
Quality: 1126.00 Length: 307  
Ratio: 4.155 Gaps: 0  
Percent Similarity: 88.274 Percent Identity: 71.336  
alignment\_block:  
US-09-898-586-24 x ABA68294/rev ..  
Align seg 1/1 to reverse of: ABA68294 from: 1 to: 933  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
|||||  
930 ATGGGGGAAATCAGACAATGCTCAGACAGAGTTCCTCTACTGGGATTTC 881  
17 ovalGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34  
|||||  
880 CTTGGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTTC 831  
34 yrvAlpheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
830 ATATCTTCACTGCTGGGGAACGGGGCCATCCTGGGCTCATCTCTACTG 781  
51 AspSerArgLeuHisAlaProMetTYrPhePheLeuSerHisLeuAlaVa 67  
|||||  
780 GACTCCAGACTCCACACCCCATGTACTCTTCTCTCTCACAACCTGGCTGT 731  
67 lValaspIleAlaTYrAlaCysAsnThrValProArgMetLeuValAsnL 84  
|||||  
730 CGTCGACATCGCCTACACCCGCAACAGGTGGCCAGATGCTGGCAACC 681

```
|||||
201 CGTCGACATCGCCTACACCCGCAACACGGTGGCCGAGATGCTGGCGAAC 250
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100
|||||
251 TCCTGCATCCAGCCAGCCATCTCTCTTGGCTGGCATGACCGACAGACC 300
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117
|||||
301 TTCTCTGTGTTGAGTTTGGACACAGCGAATGTCTCTGCTGGTCTGAT 350
117 tSerTyraSPLeuTyraValAlaIleCysHisProLeuArgTyrrLeuAlaI 134
|||||
351 GTCCTAACGATCGTTACGCTGGCCATCTGCCACCCTCTCCGATACCTCCGTC 400
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150
|||||
401 TCATGACCTGGAGAGTCTGCATCACCCCTGGCCGCTCACTTCCGACGTGT 450
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProh 167
|||||
451 GGTCCCTCTCTGCTGGCTGGCCATGTGTTCTCTCACTTCAAGACTGCCCTT 500
167 eCysArgProGlnLysIleTyrrHisPhePheCysGluIleLeuAlaVal 184
|||||
501 CTCTGGGCTTCATGAATGAAATCAACCACTTCTCTGGAATCCTGCTGTCC 550
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
|||||
551 TCAGGCTGGCTGTGCTGACACCTGGCTCAACCAAGTGTGCTCATCTTCCA 600
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217
|||||
601 GCCTGCTGTCTCTCTGCTGGGGCCACCCAGCTGTGCTGTCTCTA 650
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
|||||
651 CTGCGACATCTGGCGGGCCATCTGAGGATCCAGTCTGGGGAGGGCCGA 700
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
|||||
701 GAAAGGCTTCTCCACCTGCTCTCCACCTGCTGCTGGGACTTTC 750
251 TyrGlyThrAlaIleIleMetTyrrValGlyProArgTyrrGlyAsnProLy 267
|||||
751 TTTGGCAGTGCATCATCATGATGAGCCGCCCAAGTCCGCCATCTGA 800
267 sGluGlnLysLysTyrrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284
|||||
801 GGAGCAGCAAAAGGCTCTTTTCTATTTTTACAGTTTTTTTCAACCCACAC 850
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
|||||
851 TTAAACCCCTGATTACAGCCTGAGGAACGAGAGGATCAAGGGTGCCCTG 900
301 lysArgValLeuGlyValGlu 307
901 AGGAGAGCACTGGGCAAGGAA 921
```

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-embl/NA2001A.DAT:ABA50343

seq\_documentation\_block:  
ID ABA50343 standard; DNA; 933 BP.

XX ABA50343;  
AC  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human breast cell single exon nucleic acid probe #9038.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer; ss.  
XX

```
OS Homo sapiens.
XX
PN W0200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes
XX
PS Claim 4; SEQ ID NO 9038; 327bp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 933 BP, 243 A; 220 C; 303 G; 167 T; 0 other;
```

alignment\_scores:  
Quality: 1126.00 Length: 307  
Ratio: 4.155 Gaps: 0  
Percent Similarity: 88.274 Percent Identity: 71.336

alignment\_block:  
US-09-898-586-24 x ABA50343/rev ..

Align seg 1/1 to reverse of: ABA50343 from: 1 to: 933

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||
930 ATGGGGGAAATCAGACAATGTCACAGAGTTCTCTACTGGGATTTCT 881
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
|||||
880 CTTGGGCCCAAGATTCAGATGCTCTCTTTGGGCTTCTCTCCCTGTCT 831
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
|||||
830 ATATCTTCACCTGCTGGGGAACGGGGCCATCTCTGGGGCTCATCTCACTG 781
51 AspSerArgLeuHisAlaProMetTyrrPhePheLeuSerHisLeuAlaVa 67
```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseq-n-emb1/NA2001A.DAT:AAH323633

seq\_documentation\_block:

ID AAH32363 standard; DNA; 930 BP.

XX AAH32363;

XX 30-JUL-2001 (first entry)

XX Human olfactory receptor polynucleotide, SEQ ID NO: 936.

```

XX Human; olfactory receptor; OR; primary scent determination;
KW secondary scent determination; polypeptide library; odour receptor;
KW scent profile; scent fingerprint; scent representation; ds.
XX
OS Homo sapiens.
XX
PN WO200127158-A2.
XX
PD 19-APR-2001.
XX
PF 06-OCT-2000; 2000WO-US27582.
XX
PR 08-OCT-1999; 99US-0158615.
PR 24-FEB-2000; 2000US-0184809.
XX
PA (DIGI-) DIGISCENTS.
PA (YEDA ) YEDA RES & DEV CO LTD.
XX
PI Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;
XX
DR WPI; 2001-290713/30.
XX
PT New polynucleotides which encode polypeptides involved in olfactory
PT Sensation for identifying olfactory agonists and antagonists -
XX
PS Claim 8; Page 553; 1857pp; English.
XX
XX The present sequence is one of a number of isolated polynucleotides
CC which encode polypeptides involved in olfactory sensation. The
CC polynucleotides can be used in screening for olfactory agonists and
CC antagonists. The methods allow for the determination of primary
CC scents and the identification of the odour receptors used to detect
CC these primary scents. The methods also enable determination of
CC secondary scents and the identification of combinations of odour
CC receptors that are involved in detecting such secondary scents.
CC This enables the construction of a scent representation (also called
CC a scent fingerprint or scent profile), which may be used to re-create
CC and edit scents. Libraries of olfactory receptors are useful for
CC determining the interaction pattern of a composition with the receptors,
CC and can be used for determining differences in the olfactory faculties
CC of different individuals.
XX
SQ Sequence 930 BP; 165 A; 304 C; 219 G; 242 T; 0 other;

alignment_scores:
    quality: 1126.00      length: 307
    ratio: 4.155          gaps: 0
    percent similarity: 88.274      percent identity: 71.336

alignment_block:
US-09-898-586-24 x AAH32363 ..

Align seg 1/1 to: AAH32363 from: 1 to: 930

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17
|||||:||||| ||| :::|||||
1 ATGGGGGAAATCAGACCAATGGTCAACAGAGTCTCTCTACTGGGATTCT 50
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
:::|||||:|||||:|||||:|||||:|||||:|||||
51 CCTGGGGCCCAAGGATTCAGATGCTCCTTTGGGCTCTTCCCTGTTCT 100
34 YrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
||:::|||||:|||||:|||||:|||||:|||||:|||||
101 ATATCTTCACCTGCTGGGGAACGGGCCATCCTGGGGCTCATCTCACCTG 150
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 GACTCCAGACTCCACACCCCCCATGTACTTCTTCTCTCACACCTGGCTGT 200
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84

```



134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150  
|||||  
885 TCATGACCTGGAAGCTGCATCACTCTGGCCATCACTTCTGGACATGT 934  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuPro 167  
||| |||  
935 GGCCTCCCTCGCTATGCTCCATGTGAGCCCTCATCTTAAGACTGCCCT 984  
167 eCysArgProGlnLysIleThrHisPhePheCysGlnIleLeuAlaVal 184  
||| |||  
985 TTGTGGGCTCGTGAATGCAACCACTTCTCTGTGAATCCTGTCTGCC 1034  
184 euLysLeuAlaCysAlaAspThrHisIleasnGluAsnMetValLeuAla 200  
|||  
1035 TCAGGCTGGCTGTGTGATGACCTGGCTCAACCAAGTGTGATCATCTTGCA 1084  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
::: |||  
1085 GCCTGCATCTTCATCTGGTGGGACCACTCTGCCCTGGTGTCTCCTA 1134  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
| |||  
1135 CTCACACATCTGCGCGCATCTCGATGATCACTCTGGGAGGCGCGCA 1184  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
| |||  
1185 GAAAGGCTTCTCCACCTGCTCCCTCCACCTGCGTAGTGGAGCTCTTC 1234  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
::: |||  
1235 TTGTGCAGCGCATCTGCATGATGATGCGCCCTAAGTCCCGCATCTGA 1284  
267 sGluGlnLysLysTyrLeuLeuPheHisSerLeuPheAsnProMetL 284  
: |||  
1285 GGAGCAGCAGAGGCTCTTTTCTATTTCAGTCTTTCAACCCGATGC 1334  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||  
1335 TAAACCCCTGATTTCACAACTGAGGAATGTAGAGGTCAAGGCTGCCCTG 1384  
301 LysArgValLeu 304  
::: |||  
1385 AGGAGAGCACTG 1396

seq\_name: /SID1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS08654

seq\_documentation\_block:

ID AAS08654 standard; cDNA; 2735 BP.

AAS08654;

17-DEC-2001 (first entry)

Human cDNA encoding partial GPCR-like protein, GPCR #9.

Human; G-protein coupled receptor like protein; GPCR; immunogen;  
ophthalmic disease; neurological disease; Alzheimer's disease;  
Parkinson's disease; immunological disorder; HIV; candidiasis;  
human immunodeficiency virus; autoimmune disorder; multiple sclerosis;  
systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;  
thrombocytopaenia; aplastic anaemia; inflammatory disorder;  
septic shock; systemic inflammatory response syndrome; SIRS;  
hormonal dysfunction; cancer; atherosclerosis; wound;  
tissue regeneration; haemophilia; leukaemia; reperfusion injury;  
psoriasis; diabetes; ss.

Homo sapiens.

key location/Qualifiers  
CDS 1206..2267  
/\*tag= a  
/product= "GPCR #9"

/partial  
note="No start codon"  
FT  
XX  
XX  
PN W0200153454-A2.  
XX  
PD 26-JUL-2001.  
XX  
PE 22-DEC-2000; 2000WO-US34983.  
XX  
PR 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 20-JUN-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 31-AUG-2000; 2000US-0653450.  
PR 04-DEC-2000; 2000US-0729739.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F;  
PI Asundi V, Dermanac RT;  
XX  
DR WPI: 2001-442255/47.  
DR P-PSDB; AAU04580.  
XX  
PT New G-protein-coupled receptor-like polypeptides and polynucleotides,  
PT useful for treating diseases of ophthalmic, neurological, immunological  
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis  
PT and diabetes -  
XX  
PS Claim 1; Page 212-214; 259pp; English.

XX  
XX  
CC The sequence encodes a human G-protein coupled receptor (GPCR)-like  
CC protein. The GPCR-like polypeptides and polynucleotides are useful for  
CC the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's  
CC disease and Parkinson's disease, immunological (e.g HIV infection and  
CC candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic  
CC lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g  
CC thrombocytopaenia and aplastic anaemia), inflammatory disorders (e.g.  
CC nephritic systems. They may also be used to treat hormonal dysfunction,  
CC cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,  
CC leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples  
CC of each type of disorder are given in the specification. Anti-GPCR-like  
CC protein antibodies are useful for detecting or quantitating the  
CC polypeptide in tissue. The polypeptides can also be used as molecular  
CC weight markers and as a food supplement.  
XX

Sequence 2735 BP; 695 A; 671 C; 520 G; 849 T; 0 other;

alignment\_scores:  
Quality: 1138.00 Length: 304  
Ratio: 4.246 Gaps: 0  
Percent Similarity: 88.158 Percent Identity: 72.039

alignment\_block:

US-09-898-586-24 x AAS08654 ..

Align seg 1/1 to: AAS08654 from: 1 to: 2735

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
||| |||  
1335 ATGTGAAATATCAGACAAATGTCACAGAGTCTCTACTGGGATTTCT 1384  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34  
::: |||  
1385 CCTGGGCCCAAGATTCCAGATGCTCTCTTTGGGCTCTTCCCTGTTC 1434  
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
1435 ATGTCTTACCTGTGCTGGGAATGGGACCATCTGGGCTCATCTCACTG 1484  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67





|||||  
85 GCTCAATCCCTTATCTGTAGTCTTAGAACTCAGAAGTGAGAATACTT 36  
300 eulysArgValleuGlyValGluArgAlaIeu 310  
|||||  
35 TGAAGAGAGTGTGGAGTAGAAGGGCTTTA 4

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT.AAS08660

seq\_documentation\_block:

ID AAS08660 standard; cDNA: 2282 BP.

XX AAS08660;

XX 17-DEC-2001 (first entry)

DE Human cDNA encoding G-protein coupled receptor like protein, GPCR #13.

XX Human; G-protein coupled receptor like protein; GPCR; immunogen;  
KW ophthalmic disease; neurological disease; Alzheimer's disease;  
KW Parkinson's disease; immunological disorder; HIV; candidiasis;  
KW human immunodeficiency virus; autoimmune disorder; multiple sclerosis;  
KW systemic lupus erythematosus; rheumatoid arthritis; platelet disorder;  
KW thrombocytopaenia; aplastic anaemia; inflammatory disorder;  
KW septic shock; systemic inflammatory response syndrome; SIRS;  
KW hormonal dysfunction; cancer; atherosclerosis; wound;  
KW tissue regeneration; haemophilia; leukaemia; reperfusion injury;  
KW psoriasis; diabetes; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 1257..2189  
FT /\*tag= a  
FT /product= "GPCR #13"

PN WO200153454-A2.

PD 26-JUL-2001.

PE 22-DEC-2000; 2000WO-US34983.

XX 21-JAN-2000; 2000US-0488725.  
PR 25-APR-2000; 2000US-0552317.  
PR 20-JUN-2000; 2000US-0598042.  
PR 19-JUL-2000; 2000US-0620312.  
PR 31-AUG-2000; 2000US-0653450.  
PR 04-DEC-2000; 2000US-0729739.

XX (HYSE-) HYSEQ INC.

PI Yamazaki V, Tang YT, Liu C, Zhou P, Wang D, Zhang J, Ren F;  
PI Asundi V, Drmanac RT;

DR WPI; 2001-442255/47.  
DR P-PSDB; AAU04582.

PT New G-protein-coupled receptor-like polypeptides and polynucleotides,  
PT useful for treating diseases of ophthalmic, neurological, immunological  
PT and nephritic systems and hormonal dysfunction, cancer, atherosclerosis  
PT and diabetes -

PS Claim 1: Page 256-258; 259pp; English.

XX The sequence encodes a human G-protein coupled receptor (GPCR)-like  
CC protein. The GPCR-like polypeptides and polynucleotides are useful for  
CC the treatment of diseases of ophthalmic, neurological (e.g Alzheimer's  
CC disease and Parkinson's disease, immunological (e.g HIV infection and  
CC candidiasis), autoimmune disorders (e.g multiple sclerosis, systemic  
CC lupus erythematosus and rheumatoid arthritis), platelet disorders (e.g  
CC thrombocytopaenia and aplastic anaemia), inflammatory disorders (e.g  
CC septic shock and systemic inflammatory response syndrome, SIRS) and  
CC nephritic systems. They may also be used to treat hormonal dysfunction,

CC cancer, atherosclerosis, wounds, tissue regeneration, haemophilia,  
CC leukaemias, reperfusion injury, psoriasis and diabetes. Numerous examples  
CC of each type of disorder are given in the specification. Anti-GPCR-like  
CC protein antibodies are useful for detecting or quantitating the  
CC polypeptide in tissue. The polypeptides can also be used as molecular  
CC weight markers and as a food supplement.

XX Sequence 2282 BP; 583 A; 573 C; 428 G; 698 T; 0 other;

alignment\_scores:

Quality: 1147.00 Length: 304  
Ratio: 4.264 Gaps: 0  
Percent Similarity: 88.487 Percent Identity: 72.368

alignment\_block:

US-09-898-586-24 x AAS08660 ..

Align seg 1/1 to: AAS08660 from: 1 to: 2282

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
||| ||| ::|||  
1257 ATGTGAAAAATCAGACAAATGTCACACAGAGTTCCTCTACTGGGATTTC 1306  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
::|||  
1307 CCTGGGCCCAAGGATTGATGCTCTCTTTGGGCTCTTCTCCTGTTCT 1356  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
1357 ATGCTTTCACCCCTGCTGGGGAATGGACCATCTGGGGCTCATCTCACTG 1406  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67  
|||||  
1407 GACTCCAGACTCCACACCCCATGTACTTCTTCTCTCACACACCTGGCCGT 1456  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84  
|||||  
1457 GGTCAACATCGCCCTATGCTGCAACACAGAGTGCCTCCGATATTCATCA 1506  
84 euleuHisProAlaIysProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
1507 TCCTGCATCCAGCCCAAGCCCATCTCTTGTGCTGCATGACACAGACC 1556  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValAlme 117  
|||||  
1557 TTTCCTTTTGAAGTTTGACATACATAATGCTCTCTGTTGCTGCTGAT 1606  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
1607 GTCTACGATCGGTACGTGTCATCTGCACACCTCTCCGATATTCATCA 1656  
134 lMetThrTrpArgValCysIleThrIleAlaValThrSerTyrThrThr 150  
|||||  
1657 TCATGACCTGGAAGATCTGCATCACTCTGGCCATCACTCTCGACATGT 1706  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh 167  
|||||  
1707 GGCTCCCTGCTGCTATGTCATGTGAGCCATCTTAAGACTGCTCTT 1756  
167 eCysArgProGlnIleTyrHisPhePheCysGluIleLeuAlaVal 184  
|||||  
1757 TTGTGGCCCTCGTAATCAACCACTTCTCTGTAATCCTGTCTGTCC 1806  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
1807 TCAGGCTGGCCCTGTGTGATCACTGGCTCAACAGGTGTGATCTTTGCA 1856  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
:::|  
1857 GCTGTGATGTTTCATCTGCTGGGAGCACCTGCTGCTGCTGCTGCTTA 1906  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234

KW antifungal; vulnerary; antilucer; ss.  
XX  
OS Homo sapiens.  
XX WO200157188-A2.  
PN  
XX 09-AUG-2001.  
PD  
XX  
PF 05-FEB-2001; 2001WO-US03800.  
XX  
XX 03-FEB-2000; 2000US-0496914.  
PR 27-APR-2000; 2000US-0560875.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
OR WPI; 2001-457740/49.  
XX P-PSDB; ABB11839.  
XX  
PT Human proteins and DNA encoding sequences useful for preventing,  
PT treating or ameliorating a medical condition in a mammalian subject  
PT e.g. arthritis and cancer  
XX  
XX  
PS Claim 1; Page 753; 1963pp; English.  
XX  
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
CC invention also relates to vectors and recombinant host cells comprising a  
CC nucleotide of the invention, methods of producing the novel polypeptides,  
CC antibodies against the polypeptides, methods of detecting the nucleotides  
CC or polypeptides in a sample, and methods of identifying compounds which  
CC bind to polypeptides of the invention. Although novel, many of the  
CC polypeptides of the invention have homology to known proteins, thereby  
CC giving an insight into their probable biological activities, and hence  
CC potential therapeutic applications. The polypeptides of the invention may  
CC have various activities, including cytokine, cell proliferation or cell  
CC differentiation activities; stem cell growth factor activity;  
CC haematopoiesis regulatory activity; tissue growth activity;  
CC immunomodulatory activity; activin- or inhibin-related activities;  
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
CC thrombolytic activities; receptor or ligand activities; or may be  
CC involved in oncogenesis, cancer cell proliferation or metastasis.  
CC Depending on their biological activities, polypeptides and nucleotides of  
CC the invention are useful for preventing, treating or ameliorating medical  
CC conditions, e.g., by protein or gene therapy. Such conditions include  
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
CC vascular growth. Polypeptides involved with tissue regeneration and  
CC repair (or nucleic acids encoding them) may be used to promote wound  
CC healing (e.g., of burns, incisions and ulcers), while those with  
CC immunomodulatory activities may be used in the treatment of viral,  
CC bacterial and fungal infections in addition to immune disorders.  
CC Polypeptides with growth factor activity may be used in cell cultures to  
CC promote cell growth. For example, such polypeptides may be used to  
CC manipulate stem cells in culture to give rise to neuroepithelial cells  
CC that can be used to augment or replace cells damaged by illness,  
CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
CC may also be used in the diagnosis of the above conditions, and in drug  
CC screening techniques. The present sequence represents a cDNA encoding a  
CC novel human polypeptide of the invention.  
XX  
SQ Sequence 1315 BP; 380 A; 281 C; 360 G; 294 T; 0 other;

alignment\_scores:                      Length:                      312  
                    Quality: 1583.00  
                    Ratio: 5.106                      Gaps: 2  
Percent Similarity: 99.359              Percent Identity: 99.359

alignment\_block:

US-09-898-586-24 x ABA09073/rev ..  
Align seq 1/1 to reverse of: ABA09073 from: 1 to: 1315  
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPheP 17  
|||||  
935 ATGGGGGACATATAACATCCATCACAGAGTTCTCTACTGGGATTTC 886  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
|||||  
885 CGTTGGCCCAAGGATTCAGATGCTCCTCTTGGGCTCTTCCCTGTTCT 836  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
835 ACGTCTCACCCCTGCTGGGGAACGGACCACTAGGGGCTCATCTCACTG 786  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlay 67  
|||||  
785 GACTCCAGACTGCACGCCCCATGTACTTCTCTCTCACACCTGGGCGG 736  
67 alValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 83  
|||||  
735 TCGTCACATGCGCTACGCCCTGCACACACGTCGCCCGGATGCTGTGAAC 686  
84 LeuLeuHisProAlaLysProIleSerPheAlaGlyArgMetLeuGlnTh 100  
|||||  
685 CTCTGCATCCAGCCCAAGCCCATCTCTTGGGGCCGCATGATGCAGAC 636  
100 rPheLeuPheSerThrPheAlaValThrGlyLeuLeuLeuValAlM 117  
|||||  
635 CTTCCTGTTTTCACACTTGTGCTGTACAGAAATGTCCTCCTGCTGTTGA 586  
117 eSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAl 133  
|||||  
585 TGCTCTATGATCTGTACGTGGGCCATCTGCCACCCCTCCGATATTGGGC 536  
133 alIleMetThrTrrPargValCysIleThrLeuAlaValThrSerTrrThT 150  
|||||  
535 CATCATGACCTGGAGAGTGTGCATCACCCCTCCGGTGACTTCTGGACCA 486  
150 hrgIyValLeuLeuSerIleHisLeuValLeuLeuProLeuPro 166  
|||||  
485 CTGGAGTCTTTTATCTTGCATTCATCTGTGTACTTCTACCTTAACC 436  
167 PheCysArgProGlnLysIleTyrHisPhePheCysGlnIleLeuAlaVa 183  
|||||  
435 TTCTGTAGGCCCCAGAAATTATACACTTTTGTGAATCTTGGCGT 386  
183 IleuLysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuA 200  
|||||  
385 TCCTCAACTGCTGTGCGATACCCACATCAATGAGAACATGGCTTGG 336  
200 lagIyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSer 216  
|||||  
335 CCGGACCAATTTCTGGGCTGGTGGACCCCTTGCCACAATGTAGTTTCA 286  
217 TyrMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValG 233  
|||||  
285 TATATGTGCATCTCTGTGCTATCCTTCAGATCCAATCAAGGAAGTTCA 236  
233 nArgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuP 250  
|||||  
235 GAGGAAGCCCTTCTGCACCTGCTTCCACACCTGTGTGATTGGACTCT 186  
250 heryGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnPro 266  
|||||  
185 TTTATGGCACAGCCATTATCATGTATGTTGGACCCAGATATGGAAACCC 136  
267 LysGluGlnLysLysTyrIleLeuLeuPheHisSerLeuPheAsnProMe 283  
|||||  
135 AAGGAGCAGAAATATCTCTGCTGTTTTCACAGCCTCTTTAATCCCAT 86  
283 tLeuAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrL 300





PS Claim 8; Page 264-265; 1857pp; English.  
XX  
CC The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.  
XX  
SQ Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

alignment\_scores:  
Quality: 1584.00 Length: 310  
Ratio: 5.160 Gaps: 0  
Percent Similarity: 99.032 Percent Identity: 99.032

alignment\_block:  
US-09-898-586-24 x AAH31648 ..

Align seg 1/1 to: AAH31648 from: 1 to: 930

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17
1 ATGGGAGACAATATATACATCCATCAGAGAGTCCCTCTACTGGGATTTCC 50
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34
51 CGTTGGCCCAAGGATTCAGATGCTCCTCTTTGGGCTCTTCTCCCTGTTC 100
34 yValAlpheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
101 ACGTCTTCACCCCTGCTGGGGAACGGACCATACTGGGGCTCATCTCAGTG 150
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67
151 GACTCCAGACTGCACGCCCCCATGTACTTCTCTCTCACACCTGGCGGT 200
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84
201 CGTCGACATCGCCCTACGCCCTGCACACAGGTCGCCCGGATGCTGTGAAC 250
84 euleuHisProAlaIysProIleSerPheAlaGlyArgMetMetGlnThr 100
251 TCCTGCATCCAGCCCAAGCCCATCTCTTGGCGGCGCATGATGCAGACC 300
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValAlme 117
301 TTTCTGTCTTCCACTTTTGTGTCACAGATGTCCTCTGCTGCTGAT 350
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134
351 GTCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 400
134 lMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150
401 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCCTGACCA 450
151 GlyValLeuLeuSerIleHisLeuValLeuLeuLeuProLeuProPh 167
451 GGAGTCCTTTATCTTCATTCATCTGTGTACTTCTACCTTACCCCTT 500
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184
501 CTGTAGGCCCCAGAAAATTTATCAGCTTTTCTGTGAATCTTGCTGTTC 550
```

```
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
551 TCAACTTGCCCTGTGCAGATACCCACATCATGAGACATGCTTGCC 600
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValAlaSerTy 217
601 GGAGCAATTTCTGGGCTGTGGGACCCCTGTCCACAATTTGATTTGATA 650
217 lMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
651 TATGTGATCTCTGTGTATCTTCTCAGATCCAAATCAAGGAAGTTTCA 700
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
701 GGAAAGCTTCCGACCTGCTTCTCCACCTCTGTGTGATTGACCTCGTT 750
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267
751 TATGGCAGACCATTTATCATGTATGTTGACCCAGATATGGAAACCCCAA 800
267 sGluGlnLysLysTyrIleLeuLeuPheHisSerLeuPheAsnProMetL 284
801 GGAGCAGAGAAATATCTCTGCTGTGTTTACAGACCTCTTTAATCCATGC 850
284 eUAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
851 TCAATCCCTTATCTGTAGTCTTAGAACCTCAGAAAGTGAAGAATCTTGG 900
301 LysArgValLeuGlyValGluArgAlaLeu 310
901 AAGAGAGTGTCTGGAGTAGAAGGCTTTA 930
```

seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA509946

seq\_documentation\_block:  
ID AA509946 standard; DNA; 1040 BP.

XX AA509946;

DF 24-OCT-2001 (first entry)

XX DNA encoding human odorant receptor (OR)-like protein, NOV2.

XX  
KW Odorant receptor; human; OR; NOV2; therapeutic; neuro-olfactory system;  
KW trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;  
KW lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;  
KW acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;  
KW Albright hereditary osteodystrophy; diagnostic; ds.

XX Homo sapiens.

XX OS  
FH Key Location/Qualifiers  
FT 5'UTR 1..81

FT CDS /\*tag= a  
82..1014

FT 3'UTR /\*tag= b  
1015..1040

FT /\*tag= c

PN WO200151632-A2.

PD 19-JUL-2001.

XX 16-JAN-2001; 2001WO-US01513.

XX 13-JAN-2000; 2000US-0175989.

PR 14-JAN-2000; 2000US-0176134.

PR 25-JAN-2000; 2000US-0177839.

PR 26-JAN-2000; 2000US-0178191.

PR 26-JAN-2000; 2000US-0178227.

PR 14-JUL-2000; 2000US-0218324.



XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX  
PS Claim 8, Page 251; 1857pp; English.  
XX

CC The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.

XX Sequence 930 BP; 193 A; 265 C; 202 G; 270 T; 0 other;

alignment\_scores:                      Quality: 1584.00                      Length: 310  
  Ratio: 5.160                      Gaps: 0  
Percent Similarity: 99.032              Percent Identity: 99.032

alignment\_block:  
US-09-898-586-24 x AAH31617                      ..  
Align seg 1/1 to: AAH31617 from: 1 to: 930

```
1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPheP 17
1 ATGGAGACAAATATACATCCATCAGAGAGTCCCTCTACTCGGATTCC 50
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34
51 CGTTGGCCCAAGGATTCAGATGCTCTCTTGGGCTCTCTCCCTGTCT 100
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50
101 ACGTCTTCACTGCTGGGAGCGGACCATACTGGGCTCATCTCACTG 150
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAla 67
151 GACTCCAGACTGCACGCCCATGTACTTCTTCCCTGCACACCTGGCGT 200
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84
201 CGTGCACATCGCTTACGCTGCACACGGTGCCCGGATGCTGTGAACC 250
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetGlnThr 100
251 TCCTGCATCCAGCCCAAGCCCATCTCTTGGGGCCGATGATGCAGACC 300
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValAlaMe 117
301 TTCTGTGTTCCACTTCTTGTCTGTCAACAATGTCCTCTGCTGGTGAT 350
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 134
351 GTCTATGATCTGTACGTGGCCATGTGCCACCCCTCCGATATTGGCCA 400
134 leMetThrTrrPargValCysIleThrLeuAlaValThrSerTrrThr 150
401 TCATGACCTGGAGAGTCTGCATCACCTTCGGGTGACTTCTGGACCACT 450
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuPro 167
451 GGAGTCCTTTATCTTATCTTGATTCATCTGTGTACTTCTTACCTTTA 500
```

```
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184
501 CTGTAGGCCCCAGAAATTTATCACTTTTGTGTAATCTTGCGCTGTC 550
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200
551 TCAACTGCTGCTGTGCAGATACCCACATCATGACATGATGCTTGCC 600
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerTyr 217
601 GGAGCAATTTCTGGGCTGGTGGGACCTTGTCCACAATTGATTGATA 650
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234
651 TATGTGCATCTCTGTGCTATCTCTCAGATCCAAATCAAGGAAGTTCA 700
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250
701 GGAAAGCCTTCCGACACCTGCTTCCACCTCTGTGTGATTGACTCGTT 750
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnPro 267
751 TATGGCACAGCATTTATCATGTATGTTGGACCCAGATATGGNAACCC 800
267 sGluGlnLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMet 284
801 GGAGCAGAGAAATATCTCTGCTGTGTTCACAGCCCTCTTAATCCCA 850
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300
851 TCAATCCCTTATCTGTAGTCTTAGGAACCTCAGAAAGTGAAGAATAC 900
301 LysArgValLeuGlyValGluArgAlaLeu 310
901 AAGAGAGTGTGGAGTAGAAGGCGCTTTA 930
```

seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH31648

seq\_documentation\_block:

ID AAH31648 standard; DNA; 930 BP.

XX AAH31648;  
XX AC  
XX 30-JUL-2001 (first entry)  
XX

XX Human olfactory receptor polynucleotide, SEQ ID NO: 221.

XX Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KM scent profile; scent fingerprint; scent representation; ds.  
XX

XX Homo sapiens.

XX WO200127158-A2.

XX 19-APR-2001.

XX 06-OCT-2000; 2000WO-US27582.

XX 08-OCT-1999; 99US-0158615.

XX 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

XX (YEDA ) YEDA RES & DEV CO LTD.

XX Bellenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX WPI; 2001-290713/30.

XX New polynucleotides which encode polypeptides involved in olfactory  
PT sensation for identifying olfactory agonists and antagonists -  
XX

XX  
PI Belenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX  
DR WPI; 2001-290713/30.  
XX  
PT New polynucleotides which encode polypeptides involved in olfactory  
XX sensation for identifying olfactory agonists and antagonists -  
PS Claim 8; page 554; 1857pp; English.  
XX  
CC The present sequence is one of a number of isolated polynucleotides  
CC which encode polypeptides involved in olfactory sensation. The  
CC polynucleotides can be used in screening for olfactory agonists and  
CC antagonists. The methods allow for the determination of primary  
CC scents and the identification of the odour receptors used to detect  
CC these primary scents. The methods also enable determination of  
CC secondary scents and the identification of combinations of odour  
CC receptors that are involved in detecting such secondary scents.  
CC This enables the construction of a scent representation (also called  
CC a scent fingerprint or scent profile), which may be used to re-create  
CC and edit scents. Libraries of olfactory receptors are useful for  
CC determining the interaction pattern of a composition with the receptors,  
CC and can be used for determining differences in the olfactory faculties  
CC of different individuals.  
XX  
SQ Sequence 930 BP; 192 A; 266 C; 201 G; 271 T; 0 other;

alignment\_scores:  
Quality: 1602.00 Length: 310  
Ratio: 5.184 Gaps: 0  
Percent Similarity: 99.677 Percent Identity: 99.677

alignment\_block:  
US-09-898-586-24 x AAH32366 ..

Align seg 1/1 to: AAH32366 from: 1 to: 930

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuLeuGlyPhePr 17  
1 ATGGGGGACATATATACATCCATCACAGAGTTCCTCTACTGGGATTTC 50  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
51 CGTTGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCCCTGTTCT 100  
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
101 ACGTCTTCACCCCTGCTGGGGAACGGGACCACTGGGGCTCATCTCACTG 150  
51 AspSerArgLeuHisAlaProMetTyrPhePheLeuSerHisLeuAlaVa 67  
151 GACTCCAGACTGCACGCCCCCATGTACTTCTCTCTCACACCTGGCGGT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsn 84  
201 CGTGCACATGGCCTACGCTGCACACAGCGGTGCCCGGATGCTGGTGAAC 250  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
251 TCCTGCATCCAGCCAAAGCCCATCTCTTTGGGGCCGCATGATGCAGACC 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuLeuValValMe 117  
301 TTTCTGTTTCCACTTTGCTGTGCACAGAAATGTCTCTCTGTTGGTGAT 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
351 GTCCATATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTTGGCCA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTyrThrThr 150  
401 TCATGACCTGGAGAGTCTGCATCACCCCTGGCGGTGACTTCTTGACCACT 450

151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProH 167  
451 GGAGTCCCTTTATCCCTGATTCATCTTGTTACTTCTACCTTTACCCCTT 500  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
501 CTGTAGCCCCAGAAAATTTATCACTTTTGTGAAATCTTGGCTGTTC 550  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
551 TCAAACTTGCTGTGCAGATACCCACATCAATGAGACATGTGTTGGCC 600  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
601 GGAGCAATTTCTGGGCTGTGGGACCCCTGTCCACAATTTAGTTTCATA 650  
217 rMetCysIleLeuGlyAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
651 TATGTGATCCTCTGTGCTATCCTTCAGATCCAATCAAGGGAAGTTCAGA 700  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
701 GGAAGCCCTTCGACACTGCTTCCACACTCTGTGTGATTGACTCTTT 750  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
751 TATGGCACAGCCATTATCATGTATGTGTGACCCAGATATGGAAACCCCA 800  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284  
801 GGAGCAGAGAAATATCTCCCGCTGTTTCACAGCCTCTTATATCCCATGC 850  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
851 TCAATCCCCCTATCTGTAGTCTTAGGAACCTCAGAACTGAAGAATACTTTG 900  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
901 AAGAGAGTGTGGGAGTAGAAGGCGCTTTA 930  
seq\_name: /SIDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAH31617  
seq\_documentation\_block:  
ID AAH31617 standard; DNA; 930 BP.  
XX  
AC AAH31617;  
XX  
DT 30-JUL-2001 (first entry)  
XX  
DE Human olfactory receptor polynucleotide, SEQ ID NO: 190.  
XX  
KW Human; olfactory receptor; OR; primary scent determination;  
KW secondary scent determination; polypeptide library; odour receptor;  
KW scent profile; scent fingerprint; scent representation; ds.  
OS Homo sapiens.  
XX  
PN WO200127158-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 06-OCT-2000; 2000WO-US27582.  
XX  
PR 08-OCT-1999; 99US-0158615.  
XX  
PR 24-FEB-2000; 2000US-0184809.  
XX  
PA (DIGI-) DIGISCENTS.  
XX (YEDA ) YEDA RES & DEV CO LTD.  
PI Belenson J, Smith D, Lancet D, Glusman G, Fuchs T, Yanai I;  
XX WPI; 2001-290713/30.

PI Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;  
PI Spytek KA, Li L;  
XX  
DR WPI: 2001-451859/48.  
DR P-PSDB; AAU05142.  
XX  
PT New NOVX polypeptides and polynucleotides, useful for treating or  
PT preventing disorders of the neuro-olfactory system, cancer and multiple  
PT sclerosis  
XX  
PS Claim 9; Page 55; 141p; English.  
XX  
CC The sequence represents the coding sequence of human odorant receptor  
CC (OR)-like protein, NOV12. The NOV12 polypeptide, nucleic acid and  
CC antibody are useful as therapeutics, particularly in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease, which  
CC includes a pathology associated with NOV12 polypeptide. The NOV12 nucleic  
CC acid and polypeptide are especially useful in therapeutic or  
CC prophylactic applications for disorders of the neuro-olfactory system,  
CC e.g. those induced by trauma, surgery and/or neoplastic disorders. The  
CC DNA encoding the protein is useful in gene therapy for treating the  
CC above conditions. Furthermore, the nucleic acids and polypeptides are  
CC useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus  
CC cancer, immune response, acquired immunodeficiency syndrome (AIDS),  
CC asthma, Crohn's disease, multiple sclerosis or Albright hereditary  
CC osteodystrophy. These are also useful in developing powerful assay  
CC system for functional analysis of various human disorders, as well as  
CC in diagnostic applications.  
XX  
SQ Sequence 1014 BP; 219 A; 281 C; 216 G; 298 T; 0 other;

alignment\_scores:  
Quality: 1609.00 Length: 310  
Ratio: 5.190 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-898-586-24 x AAS09956 ..

Align seg 1/1 to: AAS09956 from: 1 to: 1014

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPheP 17  
|||||  
55 ATGGGGACAAATATAACATCCATCAGAGATTCTCTACTGGGATTTC 104  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPhe 34  
|||||  
105 CGTTGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTCTCCCTGTTCT 154  
34 yrValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
|||||  
155 ACGTCTTCACCCCTGCTGGGGAACGGGAACCATACTGGGGCTCATCTCACTG 204  
51 AspSerArgLeuHisAlaProMetIleThrPhePheLeuSerHisLeuAla 67  
|||||  
205 GACTCCAGACTGCAGCCCATGTTCTTCTCTCACACCTGGCGGT 254  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
|||||  
255 CGTCGACATCGCCTACGCTGCACACGGTGCCCGGATGCTGTGAACC 304  
84 euleuHisProAlaIleProIleSerPheAlaGlyArgMetMetGlnThr 100  
|||||  
305 TCCTGCATCCAGCCCAAGCCCATCTCTTTCGGGCGCATGATGCAGACC 354  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
|||||  
355 TTTCTGTTCCTCACTTTTGTCTGTACAGAAATGTCTCTCTGTTGGTGTAT 404  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
|||||  
405 GTCCATATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 454

134 leMetThrTrpArgValCysIleThrLeuAlaValIleThrSerTrpThr 150  
|||||  
455 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGGTGACTTCCTGGACCACT 504  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProph 167  
|||||  
505 GGAGTCTCTTATCTCTGATCATCTTGTTGTTACTTCTTACCTTACCCTT 554  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
|||||  
555 CTGTAGGCCCCAGAAATTTATACACTTTTGTGAAATCTTGGCTGTTTC 604  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
|||||  
605 TCAAACTTGCTGTGAGATACCCACATCAATGAGAACATGCTCTTGCC 654  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
|||||  
655 GGAGCAATTTCTGGGCTGGTGGGACCCCTGTCCACAAATGTAGTTTCATA 704  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGln 234  
|||||  
705 TATGTGCATCTCTGTGCTATCTCAGATCCAAATCAAGGAAAGTTTCA 754  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
|||||  
755 GGAAGCCCTTGTGCAACCTGCTCTCCCACTCTGTGTGATTGGACTCTTT 804  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267  
|||||  
805 TATGGCACAGCATATATCATGTATGTGGACCCAGATATGGAAACCCCA 854  
267 sGluGlnLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL 284  
|||||  
855 GGAGCAGAAGAAATATCTCTGCTGTTTTCACAGCCCTCTTAAATCCCATGC 904  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
|||||  
905 TCAATCCCTTATCTGTAGTCTTACGAACCTCAGAAGTGAAGAAATACCTTG 954  
301 LysArgValLeuGlyValGluArgAlaLeu 310  
|||||  
955 AAGAGAGTGTCTGGAGTAGAAGGCTTTTA 984

seq\_name: /SIDSL/gcgdata/hold-geneseq/geneseg-emb1/NA2001A.DAT:AAH32366

seq\_documentation\_block:

ID AAH32366 standard; DNA; 930 BP.

XX AAH32366;

DT 30-JUL-2001 (first entry)

DE Human olfactory receptor polynucleotide, SEQ ID NO: 939.

KW Human; olfactory receptor; OR; primary scent determination;

KW secondary scent determination; polypeptide library; odour receptor;

KW scent profile; scent fingerprint; scent representation; ds.

OS Homo sapiens.

XX Homo sapiens.

PN WO200127158-A2.

XX 19-APR-2001.

PF 06-OCT-2000; 2000WO-US27582.

XX 08-OCT-1999; 99US-0158615.

PR 24-FEB-2000; 2000US-0184809.

XX (DIGI-) DIGISCENTS.

PA (YEDA ) YEDA RES & DEV CO LTD.

CC acid and polypeptide are especially useful in therapeutic or  
CC prophylactic applications for disorders of the neuro-olfactory system,  
CC e.g. those induced by trauma, surgery and/or neoplastic disorders. The  
CC DNA encoding the protein is useful in gene therapy for treating the  
CC above conditions. Furthermore, the nucleic acids and polypeptides are  
CC useful in treating adenocarcinoma, lymphoma, prostate cancer, uterus  
CC cancer, immune response, acquired immunodeficiency syndrome (AIDS),  
CC asthma, Crohn's disease, multiple sclerosis or Albright hereditary  
CC osteodystrophy. These are also useful in developing powerful assay  
CC system for functional analysis of various human disorders, as well as  
CC in diagnostic applications.

XX  
SQ Sequence 1012 BP; 219 A; 281 C; 215 G; 297 T; 0 other;

alignment\_scores:

Quality: 1609.00 Length: 310  
Ratio: 5.190 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:

US-09-898-586-24 x AAS09955 ..

Align seg 1/1 to: AAS09955 from: 1 to: 1012

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
117 oValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
104 CGTTGGCCCAAGGATTCAGATGCTCTCTTTGGGCTCTTCCTCCCTGTCT 153  
34 yRValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
154 ACGTCTTCACCTGCTGGGGAACGGACCACTGCGGCTCATCTCCACTG 203  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAla 67  
204 GACTCCAGACTGCACGCCCATGTACTTCTCTCTCACACCTGGCGGT 253  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgmetLeuValAsn 84  
254 CGTCGACATCGCCTACGCTGCAACACGCGTCCCGGATGCTGGTGAAC 303  
84 euleuHisProAlaIleCysProIleSerPheAlaGlyArgmetLeuThr 100  
304 TCCTGCATCCAGCCCAAGCCCATCTCTTGGGGCGGCATGATGCAGACC 353  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
354 TTTCTGTTTCCACTTTTCTGCTGCACAGAATGTCCTCTCTGGTGTGAT 403  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAla 134  
404 GTCCCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 453  
134 lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr 150  
454 TCATGACCTGGAGAGTGTACCTCACCCCTCGCGGTGACTTCCGAGCCACT 503  
151 GlyValleuLeuSerLeuIleHisLeuValleuLeuProLeuProPh 167  
504 GGAGTCTTTTATCTTCTGATTCATCTTGTGTTACTTCTTACCTTTACCCCT 553  
167 eCyArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal 184  
554 CTGTAGGCCCCAGAAATATTATCACTTTTGTGTAATCTTGCGTGTTC 603  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValleuAla 200  
604 TCAAACTTGCCTGTGCAGATACCCACATCAATGAGAACATGTCCTTGGCC 653

201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValIleSerty 217  
654 GGAGCAATTTCTGGGCTGGTGGGACCCCTTGCCACAAATGTAGTTTCATA 703  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgIleValGlnA 234  
704 TATGTCATTCCTCTGTGCTATCTCTCAGATCCAAATCAAGGCAAGTTCAGA 753  
234 rGlyAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe 250  
754 GGAAGCCTTCTGCACCGCTTCTCCACCTCTGTGATGGAAGTCTTCTT 803  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProly 267  
804 TATGGCACAGCCATTAATCATATGTTGGACCCAGATATGGAAACCCCAA 853  
267 sGluGlnLysLysTyrLeuLeuLeuPheHisSerLeuPheAsnPrometL 284  
854 CGAGCAGAGAAATATATCTCTGCTGTTTTCACAGCCCTCTTAAATCCATGC 903  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu 300  
904 TCAATGCCCTTATCTGTAGCTTAGGAACCTCAGAAAGTGAAGAACTACTTG 953  
301 lysArgValleuGlyValGluArgAlaLeu 310  
954 AAGAGAGTGTGGAGTAGAAAGGCTTTA 983

seq\_name: /SIDs1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AAS09956

seq\_documentation\_block:

ID AAS09956 standard; DNA; 1014 BP.

XX AAS09956;  
XX  
XX  
XX 24-OCT-2001 (first entry)  
XX

DE DNA encoding human odorant receptor (OR)-like protein, NOVI2.

KW Odorant receptor; human; OR; NOVI2; therapeutic; neuro-olfactory system;  
KW trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;  
KW lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;  
KW acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;  
KW Albright hereditary osteodystrophy; diagnostic; ds.

XX Homo sapiens.

XX OS  
XX FH Location/Qualifiers  
XX key 1..54

XX 5'UTR /\*tag= a  
XX CDS 55..987  
XX FT /product= b  
XX FT 3'UTR 988..1014  
XX FT /\*tag= c

XX WO200151632-A2.

XX 19-JUL-2001.

XX 16-JAN-2001; 2001WO-US01513.

XX 13-JAN-2000; 2000US-0175989.

XX 14-JAN-2000; 2000US-0176134.

XX 25-JAN-2000; 2000US-0177839.

XX 26-JAN-2000; 2000US-0178191.

XX 26-JAN-2000; 2000US-0178227.

XX 14-JUL-2000; 2000US-0218324.

XX 24-JUL-2000; 2000US-0220253.

XX 25-JUL-2000; 2000US-0220590.

XX (CURA-) CURAGEN CORP.

XX



Ratio: 5.190 Gaps: 0  
Percent Similarity: 100.000 Percent Identity: 100.000

alignment\_block:  
US-09-898-586-24 x AAF58613 ..

Align seg 1/1 to: AAF58613 from: 1 to: 974

1 MetGlyAspAsnIleThrSerIleThrGluPheLeuLeuGlyPhePr 17  
1 ATGGGGGACAAATATAACATCCATCAGAGAGTTCCTCCTACTGGGATTTC 50  
17 oValGlyProArgIleGlnMetLeuLeuPheGlyLeuPheSerLeuPheT 34  
51 CGTTGGCCCAAGGATTCAGATGCTCCTTTGGGCTCTTCCTCCCTGTCT 100  
34 yValPheThrLeuLeuGlyAsnGlyThrIleLeuGlyLeuIleSerLeu 50  
101 ACGTCTTCACCCCTGCTGGGGAACGGGACCAATACTGGGGCTCATCTCA 150  
51 AspSerArgLeuHisAlaPrometTyrPhePheLeuSerHisLeuAlaVa 67  
151 GACTCCAGACTGCACGCCCCCAGTACTTCTTCCTCTCAGACCTGGCGGT 200  
67 lValAspIleAlaTyrAlaCysAsnThrValProArgMetLeuValAsnL 84  
201 GCTGCACATCGCCTAGCGCTGCAACAGCGTCCCGCGATGCTGGTAACC 250  
84 euleuHisProAlaLysProIleSerPheAlaGlyArgMetMetGlnThr 100  
251 TCCTGCATCCAGCCCAAGCCCATCTCCTTTGGCGGCGCATGATGCAGACC 300  
101 PheLeuPheSerThrPheAlaValThrGluCysLeuLeuValValMe 117  
301 TTTCTGTCTTCCACTTTTGCTGTACAGAAATGTCTCTCCTGGTGGTGA 350  
117 tSerTyrAspLeuTyrValAlaIleCysHisProLeuArgTyrLeuAlaI 134  
351 GTCCCTATGATCTGTACGTGGCCATCTGCCACCCCTCCGATATTGGCCA 400  
134 leMetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThrThr 150  
401 TCATGACCTGGAGAGTCTGCATCACCCCTCGCGTGACTTCTGCACCAC 450  
151 GlyValLeuLeuSerLeuIleHisLeuValLeuLeuProLeuProPh 167  
451 GGAGTCCCTTTATCTTGATTCATCTGTGTACTTCTTACCTTTACCCCT 500  
167 eCysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaValL 184  
501 CTGTAGGGCCCCAGAAATTATACACTTTTGTGAAATCTTGGCTGTTC 550  
184 eulysLeuAlaCysAlaAspThrHisIleAsnGluAsnMetValLeuAla 200  
551 TCMAACTTGCCGTGCAGATACCCACATCAATGAGAACATGCTTGGCC 600  
201 GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy 217  
601 GGAGCAATTCTCGGGGTGGGAGACCCTGTGCCACAATGTAGTTTCATA 650  
217 rMetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA 234  
651 TATGTGCATCCTCTGTGCTATCTCTTACATTCCAATCAAGGGAAGTTCA 700  
234 rglYsAlaPheCysrthCysPheSerHisLeuCysValIleGlyLeuPhe 250  
701 GGAAGCCCTTGTGCACCTGCTCTCCACCTCTGTGTGATGCACTCTTT 750  
251 TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy 267  
751 TATGGCACAGCCATTAATCATGTATGTTGGAGCCAGATATGGGAACCCCA 800  
267 sGluGlnLysLysTyrIleLeuLeuPheHisSerLeuPheAsnPrometL 284

|||||  
801 GGAGCAGAGAAATATCTCTCGCTGTTCACAGCCTCTTAATCCATGCG 850  
284 euAsnProLeuIleCysSerLeuArgAsnSerGluValIysAsnThrLeu 300  
851 TCAATCCCCCTTATCTGTAGTCTAGGAAGTGAAGAAATACTTTG 900  
301 lysArgValLeuGlyValGluArgAlaLeu 310  
901 AAGAGAGTCTGGGAGTAGAAGGCTTTTA 930  
seq\_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA509955  
seq\_documentation\_block:  
ID AA509955 standard; DNA; 1012 BP.  
AC AA509955;  
XX  
XX  
DT 24-OCT-2001 (first entry)  
XX  
DE DNA encoding human odorant receptor (OR)-like protein, NOV11.  
XX  
KW Odorant receptor; human; OR; NOV11; therapeutic; neuro-olfactory system;  
KW trauma; surgery; neoplastic disorder; gene therapy; adenocarcinoma;  
KW lymphoma; prostate cancer; uterus cancer; immune response; asthma; AIDS;  
KW acquired immunodeficiency syndrome; Crohn's disease; multiple sclerosis;  
KW Albright hereditary osteodystrophy; diagnostic; ds.  
XX  
OS Homo sapiens.  
XX  
FH key location/Qualifiers  
FT 5'UTR 1..53  
FT CDS /\*tag= a  
FT /\*tag= b 54..986  
FT /\*product= "Odorant receptor-like protein, NOV11"  
FT 3'UTR 987..1012  
FT /\*tag= c  
XX  
PN W0200151632-A2.  
XX  
PD 19-JUL-2001.  
XX  
PF 16-JAN-2001; 2001WO-US01513.  
XX  
PR 13-JAN-2000; 2000US-0175989.  
PR 14-JAN-2000; 2000US-0176134.  
PR 25-JAN-2000; 2000US-0177839.  
PR 26-JAN-2000; 2000US-0178191.  
PR 26-JAN-2000; 2000US-0178227.  
PR 14-JUL-2000; 2000US-0218324.  
PR 24-JUL-2000; 2000US-0220253.  
PR 25-JUL-2000; 2000US-0220590.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Padigaru M, Prayaga SK, Taupier RJ, Mishra V, Tchernev VT;  
PI Spytek KA, Li L;  
XX  
XX  
DR WPI; 2001-451859/48.  
DR P-PSDB; AAU05141.  
XX  
XX  
PT New NOVX polypeptides and polynucleotides, useful for treating or  
PT preventing disorders of the neuro-olfactory system, cancer and multiple  
PT sclerosis -  
XX  
XX  
PS Claim 9; Page 52; 141pp; English.  
XX  
CC The sequence represents the coding sequence of human odorant receptor  
CC (OR)-like protein, NOV11. The NOV11 polypeptide, nucleic acid and  
CC antibody are useful as therapeutics, particularly in the manufacture of a  
CC medicament for treating a syndrome associated with a human disease, which  
CC includes a pathology associated with NOV11 polypeptide. The NOV11 nucleic





630	TTTCTCTGTGTTGAGTTTGGACACAGCGAATGTCTCCCTGCTGCTGTGAT	581
117	tserTyraSpLeuTyrrValAlaIleCysHisProLeuArgTyrLeuAlaI	134
580	GTCTACGATCTGTACGTGGCCATCTGCCACCCTCCGATACCTCGTCA	531
134	lemetThrTrpArgValCysIleThrLeuAlaValThrSerTrpThr	150
530	TCATGACCTGGAGAGTCTGCATCAACCTGGCCGCTCACTTCCGTGACGTGT	481
151	GlyValLeuLeuSerLeuIleHisLeuValLeuLeuLeuProLeuProPh	167
480	GGCTCCCTCTGCTCTGTGGCCCATGTGTCTCATCTTAAGACTGCCCTT	431
167	ecysArgProGlnLysIleTyrHisPhePheCysGluIleLeuAlaVal	184
430	CTCTGGGCTCATGAATCAACCACTTCTTGTGTAAATCCTGTCTGTCC	381
184	eulysLeuAlaCysAlaAspThrHisIleasnGluAsnMetValLeuAla	200
380	TCAGGCTGGCCTGTGCTGACACACCTGGCTCAACCAGCTGTGTCATTGCA	331
201	GlyAlaIleSerGlyLeuValGlyProLeuSerThrIleValValSerTy	217
330	GCCTGCCGTGTCTTCTCTGTGGGGCCACCAGCCCTGTGCTTCTCTA	281
217	rmetCysIleLeuCysAlaIleLeuGlnIleGlnSerArgGluValGlnA	234
280	CTCGCACATCTCTGCCGGCCATCTCTGAGGATCCAGTCTGGGGAGGGCCCA	231
234	rgLysAlaPheCysThrCysPheSerHisLeuCysValIleGlyLeuPhe	250
230	GAAGAGCCCTTCTCCACCTGCTCTCCACCCTCTGCGGTGGGAGCTCTTC	181
251	TyrGlyThrAlaIleIleMetTyrValGlyProArgTyrGlyAsnProLy	267
180	TTTGCCAGTGCATCATCATGATGACATGCCCCCAAGTCCGCCATCTCTGA	131
267	scLuglnLysTyrLeuLeuLeuPheHisSerLeuPheAsnProMetL	284
130	GGAGCAGCAAAAGTCTTTTCTTAATTTCACAGTTTTCACACCCACACAC	81
284	euaSnProLeuIleCysSerLeuArgAsnSerGluValLysAsnThrLeu	300
80	TTTAAACCCCTGATTTCACGCCCTGAGAACGGAGAGGTCAAGGTGCCCTG	31
301	LysArgValLeuGlyValGlu	307
30	AGGAGAGCAGCTGGCCAAGGAA	10

